

```

1  TCGCGCCTTT CACCGGCACC TTGCGTCGGT CGCGCCGCGG GGCCTGCTCC
51  TGCCGCGCGC ACCCCGCGGG CTTCGGCTCC GGCACGGGTC GCGCCAGCT
101 TTCTTGCACC TGAGGCCGCC GGCCAGCCGC CGCCATGGGT GCCTACCTCT
151 CCCAGCCCAA CACGGTGAAG TGCTCCGGGG ACGGGGTCGG CGCCCCGCGC
201 CTGCCGCTGC CCTACGGCTT CTCCGCCATG CAAGGCTGGC GCGTCTCCAT
251 GGAGGATGCT CACAACTGTA TTCTTGAAGT GGACAGTGAG ACAGCCATGT
301 TTTCTGTCTA CGATGGACAT GGAGGGGAGG AAGTTGCCTT GTACTGTGCC
351 AAATATCTTC CTGATATCAT CAAAGATCAG AAGGCCTACA AGGAAGGCAA
401 GCTACAGAAG GCTTTAGAAG ATGCCCTTCT GGCTATTGAC GCCAAATTGA
451 CCACTGAAGA AGTCATTAAG GAGCTGGCAC AGATTGCAGG GCGACCCACT
501 GAGGATGAAG ATGAAAAAGA AAAAGTAGCT GATGAAGATG ATGGGGACCA
551 CTTCTATAAG AGAAACAAGA ACCTGCCACC TGAGGAACAG ATGATTTCAG
601 CCCTTCCTGA CATCAAGGTG CTGACTCTCA CTGACGACCA TGAATTCATG
651 GTCATTGCCT GTGATGGCAT CTGGAATGTG ATGAGCAGCC AGGAAGTTGT
701 AGATTTTCATT CAATCAAAGA TCAGCCAGCG TGATGAAAAT GGGGAGCTTC
751 GGTATTGTG ATCCATTGTG GAAGAGCTGC TGGATCAGTG CCTGGCACCA
801 GACACTTCTG GGGATGGTAC AGGGTGTGAC AACATGACCT GCATCATCAT
851 TTGCTTCAAG CCCCAGAAAC CAGCAGAGCT CCAGCCAGAG AGTGGCAAGC
901 GAAAACATAA GGAGGTGCTC TCTACTGAGG GGGCTGAAGA AAATGGCAAC
951 AGCGACAAGA AGAAGAAGGC CAAGCGAGAC TAGCAGTCAT CCAGACCCCT
1001 GCCCACCTAG ACTGTTTTCT GAGCCCTCCG GACCTGAGAC TGAGTTTTGT
1051 CTTTTTCCTT TAGCCTTAGC AGTGGGTATG AGGTGTGCAG GGGGAGCTGG
1101 GTGGCTTAC TCCGCCCAT CCAGAGAGG CTCTCCCTCC AACTGCAGC
1151 CCGGAGCCTC TGCTGTCCTT CCCAGCCGCC TCTGCTCCTC GGGCTCATCA
1201 CCGGTTCTGT GCCTGTGCTC TGTTGTGTTG GAGGGAAGGA CTGGCGGTTC
1251 TGGTTTTTAC TCTGTGAAGT TTATTTAAGG ACATTCTTTT TTATTGCGCG
1301 CTCCATGGCC CTCGGCCGCT TGCACCCGCT CTCTGTTGTA CACTTTCAAT
1351 CAACACTTTT TCAGACTAAA GGCCAAAACC TAATCGTTAA AAAAAAAAAA
1401 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA (SEQ ID NO:1)

```

FEATURES:

5' UTR: 1-134
Start: 135
Stop: 981
3' UTR: 984

Homologous proteins:

Top 10 BLAST Hits

| | Score | E |
|--|-------|-------|
| CRA 18000005093929 /dataset=nraa /length=546 /altid=gi 4505999 ... | 301 | 6e-81 |
| CRA 18000005157254 /dataset=nraa /length=542 /altid=gi 6679793 ... | 284 | 1e-75 |
| CRA 18000005062172 /dataset=nraa /length=543 /altid=gi 3122591 ... | 283 | 1e-75 |
| CRA 1000682325310 /dataset=nraa /length=356 /altid=gi 7508306 /... | 161 | 1e-38 |
| CRA 18000004879328 /dataset=nraa /length=348 /altid=gi 1352682 ... | 161 | 1e-38 |
| CRA 18000004973158 /dataset=nraa /length=370 /altid=gi 1171963 ... | 154 | 9e-37 |
| CRA 18000005006718 /dataset=nraa /length=414 /altid=gi 1171964 ... | 154 | 1e-36 |
| CRA 18000004973159 /dataset=nraa /length=414 /altid=gi 1076914 ... | 152 | 3e-36 |
| CRA 89000000194666 /dataset=nraa /length=352 /altid=gi 7291977 ... | 151 | 8e-36 |
| CRA 18000004972853 /dataset=nraa /length=464 /altid=gi 6320934 ... | 145 | 4e-34 |
| CRA 89000000203611 /dataset=nraa /length=662 /altid=gi 7302240 ... | 141 | 9e-33 |
| CRA 18000004915558 /dataset=nraa /length=468 /altid=gi 6319415 ... | 140 | 2e-32 |
| CRA 18000005051471 /dataset=nraa /length=468 /altid=gi 1622933 ... | 140 | 2e-32 |

| Year | 1970 | 1971 | 1972 | 1973 | 1974 | 1975 | 1976 | 1977 | 1978 | 1979 | 1980 | 1981 | 1982 | 1983 | 1984 | 1985 | 1986 | 1987 | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 | 2025 | 2026 | 2027 | 2028 | 2029 | 2030 | 2031 | 2032 | 2033 | 2034 | 2035 | 2036 | 2037 | 2038 | 2039 | 2040 | 2041 | 2042 | 2043 | 2044 | 2045 | 2046 | 2047 | 2048 | 2049 | 2050 | 2051 | 2052 | 2053 | 2054 | 2055 | 2056 | 2057 | 2058 | 2059 | 2060 | 2061 | 2062 | 2063 | 2064 | 2065 | 2066 | 2067 | 2068 | 2069 | 2070 | 2071 | 2072 | 2073 | 2074 | 2075 | 2076 | 2077 | 2078 | 2079 | 2080 | 2081 | 2082 | 2083 | 2084 | 2085 | 2086 | 2087 | 2088 | 2089 | 2090 | 2091 | 2092 | 2093 | 2094 | 2095 | 2096 | 2097 | 2098 | 2099 | 2100 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1970 | 1971 | 1972 | 1973 | 1974 | 1975 | 1976 | 1977 | 1978 | 1979 | 1980 | 1981 | 1982 | 1983 | 1984 | 1985 | 1986 | 1987 | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 | 2025 | 2026 | 2027 | 2028 | 2029 | 2030 | 2031 | 2032 | 2033 | 2034 | 2035 | 2036 | 2037 | 2038 | 2039 | 2040 | 2041 | 2042 | 2043 | 2044 | 2045 | 2046 | 2047 | 2048 | 2049 | 2050 | 2051 | 2052 | 2053 | 2054 | 2055 | 2056 | 2057 | 2058 | 2059 | 2060 | 2061 | 2062 | 2063 | 2064 | 2065 | 2066 | 2067 | 2068 | 2069 | 2070 | 2071 | 2072 | 2073 | 2074 | 2075 | 2076 | 2077 | 2078 | 2079 | 2080 | 2081 | 2082 | 2083 | 2084 | 2085 | 2086 | 2087 | 2088 | 2089 | 2090 | 2091 | 2092 | 2093 | 2094 | 2095 | 2096 | 2097 | 2098 | 2099 | 2100 | |

Expression information from BLAST dbEST hits:

Expression information from PCR-based tissue screening panels:
Human leukocyte

```

1  MGAYLSQPNT VKCSGDGVGA PRLPLPYGFS AMQGWRVSME DAHNCIPELD
51 SETAMFSVYD GHGGEVALY CAKYLPIIK DQKAYKEGKL QKALEDAFLA
101 IDAKLTTEEV IKELAQIAGR PTEDEDEKEK VADEDDGDHF YKRKNLPPPE
151 EQMISALPDI KVLTLTDDHE FMVIACDGIW NVMSSQEVVD FIQSKISQRD
201 ENGELRLLSS IVEELLDQCL APDTSGDGTG CDNMTCIIC FKPRNTAELQ
251 PESGKRKLEE VLSTEGAEEN GNSDKKKKAK RD (SEQ ID NO:2)

```

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

233-236 NMTC

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 4

```

1      10-12 TVK
2      197-199 SQR
3      253-255 SGK
4      273-275 SDK

```

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 10

```

1      38-41 SMED
2      57-60 SVYD
3      106-109 TTEE
4      122-125 TEDE
5      164-167 TLTD
6      184-187 SSQE
7      197-200 SQRD
8      210-213 SIVE
9      224-227 TSGD
10     229-232 TGCD

```

[4] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 2

```

1      2-7 GAYLSQ
2      266-271 GAEENG

```

[5] PDOC00009 PS00009 AMIDATION
Amidation site

253-256 SGKR

[6] PDOC00792 PS01032 PP2C
Protein phosphatase 2C signature

55-63 MFSVYDGHG

BLAST Alignment to Top Hit:

>CRA|18000005093929 /dataset=nraa /length=546 /altid=gi|4505999
/def=ref|NP_002698.1| protein phosphatase 1G (formerly
2C), magnesium-dependent, gamma isoform; protein
phosphatase 1G (formerly 2C);; protein phosphatase 2,
catalytic subunit, gamma isoform [Homo sapiens]
/org=Homo sapiens /taxon=9606
Length = 546

Score = 301 bits (763), Expect = 6e-81
Identities = 146/146 (100%), Positives = 146/146 (100%)

Query: 137 GDHIFYKRKNLPPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI 196
GDHIFYKRKNLPPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI
Sbjct: 401 GDHIFYKRKNLPPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI 460

Query: 197 SQRDENGELRLLSSIVEELLDQCLAPDTSBGDTGCDNMTCTIIICFKPRNTAELQPESGKR 256
SQRDENGELRLLSSIVEELLDQCLAPDTSBGDTGCDNMTCTIIICFKPRNTAELQPESGKR
Sbjct: 461 SQRDENGELRLLSSIVEELLDQCLAPDTSBGDTGCDNMTCTIIICFKPRNTAELQPESGKR 520

Query: 257 KLEEVLTSTEGAEENGNSDKKKKAKRD 282
KLEEVLTSTEGAEENGNSDKKKKAKRD
Sbjct: 521 KLEEVLTSTEGAEENGNSDKKKKAKRD 546 (SEQ ID NO:4)

Score = 284 bits (718), Expect = 1e-75
Identities = 137/139 (98%), Positives = 138/139 (98%)

Query: 1 MGAYLSQPNTVKCSGDBGVAPRLPLPYGFSAMQGWRVSMEDAHNCIPELDSETAMFSVYD 60
MGAYLSQPNTVKCSGDBGVAPRLPLPYGFSAMQGWRVSMEDAHNCIPELDSETAMFSVYD
Sbjct: 1 MGAYLSQPNTVKCSGDBGVAPRLPLPYGFSAMQGWRVSMEDAHNCIPELDSETAMFSVYD 60

Query: 61 GHGGEVALYCAKYLDPDIKDQKAYKEGKLQKALEDAFLAIDAKLTTEEVIKELAQIAGR 120
GHGGEVALYCAKYLDPDIKDQKAYKEGKLQKALEDAFLAIDAKLTTEEVIKELAQIAGR
Sbjct: 61 GHGGEVALYCAKYLDPDIKDQKAYKEGKLQKALEDAFLAIDAKLTTEEVIKELAQIAGR 120

Query: 121 PTEDEDEKEKVADEDDGDH 139
PTEDEDEKEKVADEDD D+
Sbjct: 121 PTEDEDEKEKVADEDDVDN 139 (SEQ ID NO:5)

>CRA|18000005157254 /dataset=nraa /length=542 /altid=gi|6679793
/def=ref|NP_032040.1| fibroblast growth factor inducible
13 [Mus musculus] /org=Mus musculus /taxon=10090
Length = 542

Score = 284 bits (718), Expect = 1e-75
Identities = 139/146 (95%), Positives = 141/146 (96%), Gaps = 1/146 (0%)

Query: 137 GDHIFYKRKNLPPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI 196
GDHIFYKRKNLPPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI
Sbjct: 398 GDHIFYKRKNLPPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI 457

Query: 197 SQRDENGELRLLSSIVEELLDQCLAPDTSBGDTGCDNMTCTIIICFKPRNTAELQPESGKR 256
SQRDENGELRLLSSIVEELLDQCLAPDTSBGDTGCDNMTCTIIICFKPRNT ELQ ESGKR
Sbjct: 458 SQRDENGELRLLSSIVEELLDQCLAPDTSBGDTGCDNMTCTIIICFKPRNTVELQAESGKR 517

Query: 257 KLEEVNSTEGAEENGNSDKKKKAKRD 282
 KLEE LSTEGAE+ GNSD KKKAKRD
 Sbjct: 518 KLEELSTEGAEDTGNSD-KKKAKRD 542 (SEQ ID NO:6)

Score = 279 bits (706), Expect = 3e-74
 Identities = 133/139 (95%), Positives = 138/139 (98%)

Query: 1 MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPELDSETAMFSVYD 60
 MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPELD+ETAMFSVYD
 Sbjct: 1 MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPELDNETAMFSVYD 60

Query: 61 GHGGEVALYCAKYLPDIIKDQKAYKEGKLQKALEDAFLAIDAKLTTEEVIKELAQIAGR 120
 GHGGEVALYCAKYLPDIIKDQKAYKEGKLQKAL+DAFLAIDAKLTTEEVIKELAQIAGR
 Sbjct: 61 GHGGEVALYCAKYLPDIIKDQKAYKEGKLQKALQDAFLAIDAKLTTEEVIKELAQIAGR 120

Query: 121 PTEDEDEKEKVADEDDGDH 139
 PTEDED+K+KVADEDD D+
 Sbjct: 121 PTEDEDDKDKVADEDDVDN 139 (SEQ ID NO:7)

Hammer search results (Pfam):

| Model | Description | Score | E-value | N |
|---------|------------------------|-------|---------|---|
| PF00481 | Protein phosphatase 2C | 176.2 | 4.2e-49 | 2 |
| PF01722 | BolA-like protein | 3.7 | 9.5 | 1 |

Parsed for domains:

| Model | Domain | seq-f | seq-t | hmm-f | hmm-t | score | E-value |
|---------|--------|-------|--------|-------|----------|-------|---------|
| PF00481 | 1/2 | 25 | 102 .. | 1 | 95 [. .] | 74.1 | 4.6e-20 |
| PF01722 | 1/1 | 99 | 110 .. | 76 | 87 .] | 3.7 | 9.5 |
| PF00481 | 2/2 | 137 | 228 .. | 197 | 301 .] | 102.2 | 4.5e-28 |

```

1 AAAGAATCTT TTTTTTTTTT TTGAGACGGA GTTGCTCTGT CACCCAGGGT
51 GGAGTGCAGT GGCGCCATCT TGGTTCACCT CAACCTCCGC CTCTGGGGTT
101 CAAGCTATTG GCCTGCCTTA GCCTCCCAAG TAGCTGGGAT TACAGGAGCG
151 CACCAGTACG CCTGGCTAAT TTTTGTATTT TTAGTAGAGA CGGGTTTCAC
201 ATGTTGGCCA GGCTGGTCTC GAACTTCTGG CCTCAAGTGA TCCACCACCC
251 CCCCTTGGCC TCCCAAAGTG CTGGGATTAC AAGTGTGAGC CACTGTGCCC
301 GGCTGAAAAG AATCAATTTT GTCATAGTTT GGAGAATTTT TCCTTTTCTC
351 TCCATCCCTT GAATGCAATT TATTACCAA TCTGTCTTAT TTGTTATTGT
401 CTAATTTGTC CTTTCATCTG GATTCCCAT GCCACCCTGC GTGGTACCAC
451 CTTACTCCCA GTCCTTCTCA TCTCCTGCTT AGAGTAGAG CTCTCTAACT
501 AGTAGCAGTG CCCCAGGCCA GGCGCGGTGG CTCACGCCTG TAATCCCAGC
551 ACTTTGGGAG GCTGAGGCGG GTAGATCACG AGGTGAGGAG TTCGAGACCA
601 GCCTGGCCAA CATGGTGAAA CCCCCTCTCT ACTAAAAATA CAAAAATTAG
651 CCAGGCGTGG AATCCTAGCT ACTCGGAGG CTGAGGCAGG AGAATCGCTT
701 GAACCTGGGA GGCGGAGGTT GCAGTGAGCC AAGATAGCGC CACTGCACTG
751 CAGCCTGGGC AACAAGAGCG AAACCTCTGTC TTAACAAAAA AACAATAGTA
801 GGCGGGGTGC GGTGGCTCAC GCCTGTAATC CCAACACTTT GGGAGGCCGA
851 GGCGGGCGGA TCACGAGGTC AGGAGATGGA GACCATCCTG GCTAACACGG
901 TGAAACCTCG TCTCTACTAA AAATACAAAA AATTAGCCAG GCGTGGTGGC
951 GGGCGCCTGT AGTCCAGCT ACTCGGAGG CTGAGGCAGG AGAATGGCGT
1001 GAACCCGGGA GGCGGAGCTT GTAGCCTGGG CGACAGAGCG AGACTCTGTC
1051 TCAAAAACAA ACAAAAAACA ACAACACAAC AGTGCCCCAG ACTCTCTCCC
1101 TCCAATGTAC ACTGCATACA AAGACTAGAC AAACGATGCC AAAGGTTACA
1151 CCAGGACGAC AATGAAGTCC AAGTCACTCA TCTTGGCGCT TCTTTTATC
1201 AAGCTAGCTA GTTAATATTA CCACTTACAA TCATTCTCCG AGTCCAGCCA
1251 GCTTCCTCAG AAATCCTCCC CAAATGCAGT TCACATCTT CCCTCACTCC
1301 TCCAAACCTT ATAACATTAT TCCCTTTTCC TGTGGCATAA TGCAATCCAG
1351 AGGCATCCTC TCTGAGAGCC TACATCCTTT CAATCCTCCA AGAAACAGCT
1401 CCTCCTCCCA TCCTTGAGCT TTTCCACCC AGAATAGGCT GTACCAAACA
1451 TTTCTACATT GTATTAACTC GAATTATATC ACAGATCGAG GTCTTCGCTT
1501 TCAGAAAAGA CTCACATTCT TCTCATATAG GCCTCGCATG GCCTGGCCAA
1551 GTTCTCCTA AGCTCGGAAC AAATACTGGT CAACTTTAAT TGAACCAAAT
1601 CGGGCGGGGT TTGCGGAGTC TGAGAGTGCA ACGTTGGGGA GAGGGGGATG
1651 AAAACACTGG AGGACGGGCT GAAAGCGTCG AGTCCGACAC AAAAGAGGCG
1701 TCAGACAAAA CGCCAAGAGG CTGGGGACTG GGAACGAAGG AAGAGGTTCT
1751 GCCAGAGGCG ACCTGCCACC TGCGCGAGGA AGCGGAGTAG GACGGCGGCC
1801 GTTGGTGGGC GTGGTCGCGC TAGTCTCGCG GGAGCGCGCG TTGGGCGGGC
1851 CGTTGTCCCC TGCGGGCGGG GCGAGTTGCT AAGGAAATGA CTGCCCCGAG
1901 CGCCTGGCCC CGCCGCGCAG GCCGGGCGGG GTCTGGAGCG GCGCCGTTTC
1951 CGCTTCCGCT CCCTCAGAGC TCCCGTCCCG TTACCGCCTC CTGGCCGCGC
2001 TCGCGCCTTT CACCGGCACC TTGCGTCGGT CGCGCCGCGG GGCCTGCTCC
2051 TGCCGCGCGC ACCCCCGGGG CTTGCGCTCC GGCACGGGTC GCGCCAGCT
2101 TTTCTGCACC TGAGCCGCCG GCCAGCCGCC GCCATGGGTG CCTACCTCTC
2151 CCAGCCCAAC ACGGTGAAGT GCTCCGGGGA CGGGGTGCGC GCCCCGCGCC
2201 TGCCGCTGCC CTACGGCTTC TCCGCCATGC AAGGCTGGCG CGTCTCCATG
2251 GAGGTGAGGA GGCAGGGGCC CATAGGCTGG CCGCTGCGGG GCGGGAATCT
2301 GACGGAGAAA GAGAGCGGGG GATGGGTGCT TCCCCTGGGA AGGGTCCCAA
2351 TTGGGAGCCT GCGGCCGCAG CGGCCGTTTG CGGGGCGACA GAGACCGCGG
2401 GGTCAGGGCC GAGAGGGAGC TCCCGATGCT TGGGACCGT GTGCCGGTGG
2451 CCAGTGGCGG CGAGGGCTCG GCCATGTGGG AAGAGGCACC TTCCGCCAC
2501 TGACCGCCTT CTCCCCGAGC TTTGGCGCCA TCCTCTCGTG CCAACCTAGC
2551 CCTCCAGGCT CATCAACGGT GTGAGGTTTA GTGTGGGAGT AAAGACGCAA
2601 ATAGGGGCTA TTTATTCATT TGTTTTTCAA GGGAGAGGTG AATAGATGTG
2651 AATAACTTTT TAAATTTTAA TATTTAAAAT ATCTGATGTG GGAAGCCTCT
2701 TTTGGCTAGG AGTTTGACAG TGAAAGGAAC CCCGGGCAGA GTCTGTTTCA
2751 CATTTTGGTT GCCTGGCCTT GGGCTCTTGC ATGTTAATTT CAGAGGCTGG
2801 ACCCGACCTC CAGGAGTTGT CACTCATTTG CACTCTTTTC AGGGCCTTTT
2851 ACTAACTTCG GAAAACCTGA ATTATGTCAG TCCCTAGGTT TTCCTTTTTA
2901 TTATGCTTCG TTTTCTTTCT CTCTTAGGAT TTCTCTAAAA CTTAATCAGT
2951 AATTCTCATT TGCTGTAAAT TGTAAGTTAT AGTTTCTTTG ATTGTAAGTC
3001 ATTTTCATCA TTTTTTTCAT CACAACCTAC CTACAAAGGG CTTTCTTAGA
3051 AAATTTTACT CTGGACAAAA GGGGAAAAGA AAAATATTGG GGGGAAAGTA
3101 GTAGTATTAG GTAAACTTGT ATGTGAAACT ACAAAGAGA AGAGGGAAAA

```

FIGURE 3, page 1 of 17

```

3151 CTGCGGTAGG GAGGAAAGGG AGGAAGACGG GTTAACCGTG GCTTTGTGAA
3201 GAGCATTCTG AAGTCTAGGC AAAAGGGCCA GGGAAATACT CTGTCTGGTA
3251 TTGAGGGTTT CTCCACCTAC CGGGTGGGCT TCAGGTAACA GCGAAATACT
3301 GTCTCCCTTG GGAATTGTTT CAGATCCCTC GCTCCTCCTG TGGTTAGCTC
3351 TGAATGCCA GTATGAACCT CAATGTTTTG TTTTCCGATT CAAATTTTAT
3401 ATTCATAACT GACCTTAATA ACAATTTTAC AATTAGGTAT AAAATTTTCA
3451 GATCCTAGTG TATCCTATAG TTCATCTCAT CTGCTTTGGC TCCCTTTTTT
3501 TTTTTTTGAG ACAGAATTC GCTCTGTGTG CCCAGGCTGG AATGCAATGG
3551 CGCGATCTCG GCTCACC GCA ACCTCCGCT CCCAGGTTGG AGCAATTCTC
3601 CTGCCTCAGC CTCCATAGTA GTTGGGATCA CAGGCATGTG CCACCACGCG
3651 TGGCTTATTT TGTATTTTTA GTAGAGACAG GGTTCCTTCA TGTGGGTGAG
3701 GCTGGTGTG AACTCCTGAC TTAGGTGATC TGCCTGCCTC GGTCTCCCAA
3751 AGTGCTACCA CCACNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
6001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
6051 NNNNNNNNNN NNNNNNNNNN NNNNNNTGGA TTACAGACAT GAGCCACTGC
6101 GCCAGCCTT ATTTAGAAAT TTCTTAGTG AAAGATGATA AATTTTCAGT
6151 TTTTCATTAT CTGAACATGT TTTTATCTAG CCTTTGTTCT GAAAAGATGC
6201 TTGACTCAG TACCCAGTTC TAGATTGACA GTTAATTTT CTTAATTTGT
6251 AAATGTTGTT TCATTGATTG ACTTCCATTG TTGTTGCGAA AAATTTATCA

```

FIGURE 3, page 2 of 17


```

6301 TCAGCCATTT CTGACTTTTG ATCTGTGTTT TCTCTTTGGT TTCTCTTTTT
6351 TTTTCTCTTT TTTTCTTTTT GAGACGGAGT GTCGCTCTGT TGCCCAGGCT
6401 GGATGCGAGT GGCATGATCT TGGCTCACTG CAACCTCTGC CTCCTGAGTT
6451 CAAGCGATTG TCCTGCCTCA GCTTCCCGAG TAGCTGGGAT TACAGGCGCC
6501 TGCCACCATG CCCGGCTAAT TTTTGTATT TTCAATAGAG ACAGGGTTTC
6551 ACTATGTTGG CCGGGTGGT CTCCAACCTC TGACCTCTTA ATCCGCCCGC
6601 CTCGGGCTCC CAAAGTGCTG GCATTACAGG CGTGAGCCAC CATGCCTGGC
6651 CCATCAGTTG ATGTAGTCTT AAGGGGACAA GAGTACATTT AATATTTGGT
6701 TGTGAGGTTT TCTGGAAGTG ACAAACCTGC TTTCTATGGA GAGTTAGGAA
6751 TTTTCTTTTT TTTTGAAACG GAGTCTCGCA TTGTCACCCG GGCTGGAGTG
6801 TAGTGGCTTG ATCTCGGCTC ACTGAACTT CCGCCTCTTG GGTTCAGTG
6851 ATTCTCTGCT CTCAGCCTTC CAAATAGCTG GGATTACAGG AGTCTGCCAC
6901 CAGGCCAGCT AATTTTTTTT TGTATTTTGA GTGGAGACAG GATTCTACTA
6951 TGTTGGCCAG GCTGGTCTCA AGACTCCTGA CGTTGTGATC CACCTGCCTC
7001 GGCTTCCCAA AGTGCTGGGA TTACAGATGT GAGCTACCGT GCCCGGCCAG
7051 GAATTTTTTG TGCTATAAAT CATATTTTCC TTTTATTAAA GGCAGTGTC
7101 ATATCTATAG TATAATTTTG AGGAGGCTGG CTATTTATTG CTGTGTAGAA
7151 GCTGGCTTAT TAGTGGTCAA GGGGTCATCT AGAATTGACT ATAAAGATAG
7201 TATTGAGCAG AAAATCTCTA AAATACCTGC ATATTAGTTT CAGTCATTAA
7251 ATTAATGGAA AAAATATAAA AAGAAATATC ACAAGTATGC TATGGGGTTC
7301 TACCTTAGGG CATTGAAGGT TGAAAAACAT TTATTTTCTT ATCTTCAGAT
7351 TAGCATCTCA TATCAATAGC CAATAGCTTA AAGCGCTTTT TACTTACTAA
7401 ACCAGGTCAG AATTTCTCTC TCTCTCTCTT TTTTTTTTTT AGACAGAGTC
7451 TCTGTACCCC AGTCTAGAGT GCAGTGGTGT GATCTCGGCT CACTGGCCTT
7501 TCTGTTTAAAG TGATTCTCCG GCCTCAGCCT CCTGAGTAGC TGGGATTACA
7551 GGCATGTGCC ACGACACTTG GCTTATTTT TGTATTTTGA GTAGAGATGG
7601 GGTTTCTCTG TGTTGGTCAG GCAGGTCTTG AACTCCTGAC CTCAGGTGAT
7651 CTGCCCCGCT CCGCCTCCCA AAGTGCTGGG ATTACAGACG TGAGCTACTG
7701 CGCCTGGCCA GAATTTCTTT GTCTAGAATG TGGTTAGCAA CTTTTATAAA
7751 AACGCATTAT TTGCATTGA TTAGCATGCA GTACCCATTC ACAGTTCAAA
7801 GCTAGTATAG AATTATATCA CATGTATGCC CATGAGCATG GAGAACTAT
7851 TTTCTTTTGA TTTTCTTAAAG TTGGAGTTT GCTCTTGTG CCCAGGCTGG
7901 AGTGCAATGG TGCCATCTCG GCTCACTGCA GCTTCTGCCT CCTGGGTCT
7951 AGCAATTTGC CTACCCAGC CTCCCAAGTA GCTGGGATTA CAGGCACTCG
8001 CCACCATGCC CAGCTAACTT TTTTGTATTT TTAGTAGAGA AGGGGTTTTC
8051 CCATGCTGGC GAGTCTGGTC TTGAACCTCT GGCTCAAGT GATCTGCCCG
8101 CCTCAGCCTC CCAAAGTGCT GGAATTACAG GCATGAGCCA CTGTGCCCGG
8151 CCTTTTATT TTTTAAATTA TTTATGTATT TATTTTGAGA CAGGATCTCA
8201 CTCTTGCCCA TGCTTGAGTG GTAAGGAGTA TGGGATTCTT TGTGCCGTT
8251 CCTTCACATC CTCACACAC TTATCTGCCT TTACAGTGGC TCGATCATAG
8301 TTCCTGTCAT AGCCTTCTGG GCTCAAGGGG TCTTCCAGCC TCAGCCTAAT
8351 ATAGGCACAT GCCACCATGA CTGGCTAATT TTTTTTTTTT AAGTTTTTTT
8401 TTGTAGAGAT AGGGCCTTGC AGTGTGCCC AGGCTGAGGA ATTTTATTTA
8451 TGTTTATTTT ATTTATTTAT TTATTTATTT ATTTATTTAT TTATTTATTT
8501 ATTGAGATGG AGTCTTACAC TGTCACCCAG GCTGGAGTGC AGTTGCGCGG
8551 TCTCAGCTAC TGCAAGCTCC GCCTCCCGGG TTCATGCCAT TCTCCTGCCT
8601 CAGCCTCCCG CGTAGCTGGA CTACAGGCGC CCGCCACCAC GCCTGGCTAA
8651 TTTTTTGTGT GTGTTTTTAG TAGAGGCGGG GTTTCACCAT GTTAGCCAGG
8701 ATGGTCTCGA TTCCTGACCT CGTGATCCAC CCACATCGGC CTCCCAAAGT
8751 GCTGGGATTA CAGGTGTGAG CCACCATGCC TGGTCTAGAA ATTATTTTAT
8801 ATTTTATACC ATTGCCTTAT AAGTTCTCAA GCAACTGGAA AATACAATCA
8851 GAACGTATTC CTCAAGATT CAAGGATATT TTACACAAAG TTCTATTGTC
8901 TGATTCCTTA GCAGTTGTTA CTACTGTTTC CCTAACCTCT AATCTTCTAT
8951 TGGGTTATTA GTCTTAGAAT TGAATTTTGA GAGGTAAGGG CTGGAATTTG
9001 AACATAGAAA TTTATACAGG TCTGATCAGT AGTTCCTGAC ATTGTATTAT
9051 CTGGAACAA ATCTTTAGAA CTGAGCTTAA GATGTTTAAAT GACATTTTGT
9101 AGACAGAGTA TGATTTTCACT GTAGTTGTTT TTGTTTCTTT CTAGATCTAG
9151 TTCAGAGATG AAGTATATCA ACTTTTTTTT CCTTTTGTAC CCAATGCTAG
9201 CAGAAAAACA ACACCTTTTA ATCATATTTA GTATTTGAAA ATGTGTATAC
9251 AGGTTCCCTT TTATTTTATT TATTTCTTAC AGGTTCCCTT TTAATCAGCT
9301 TTATTGAGAT AGAGTTCATA TACTGTATGG TTCATACCAC ATATGGTTCA
9351 TATACCATAC AGTATATGAA CTCACCTTAA GAGTATAATT CAGTGGGTTT
9401 AAGGGTATAA TTCATTCAAT TTAAGGGTAT AATTCAGTGG CTTTATAGTAT

```

FIGURE 3, page 3 of 17

9451 ATTTTCTTTT TTTTCTTCTT TTTTCTTTT TTCGAGACAG GGTCTTGCTC
9501 TGTACCCAG GCTGGAGTGC AGTGACGCAG CCTCAGCTCA CTGCAACCTC
9551 CACCTCCCAG GTTCAAGCGA ATTCTCTTGC CTCAGCATCC TGAGTAGCTG
9601 GAACTACAGG CTCACGCCAC CACACCCAGC TAATTTTAT ATTTTCAGTA
9651 GAAACAGGGT TTCACCATGT TGCCCAGGCT GGTCTCGAAC TCCTGGCCTC
9701 AAGTGATCTG CCTGCCTCAG CCTCCCAAAG TGCTGGGATT ACAGGCATGA
9751 GCCACCGTGC CTGCCTGTTT TGTAAGTGTAT TCAAACAGAG TTGTACAAC
9801 GTCACCACAA TCAGTTTATG AACCCCAAAA AGAAACCCTG TACTCTTTAC
9851 CAGTCACTCC CTATCTTCCG TCCACTAACC CCTGGCATCC ACTAATTAC
9901 ATGACCTCTA TGAATTTGCC TATTCTGAAC ATTTTATAAA TGGAATTCTA
9951 AATACACTAC CTTTATATATC TGGATGCTTT TACTAAGCAT GTGTATTTT
10001 GAAATTGACT TTAAAGCTTG TTGGCCCTG GAAGAGTAA TTAATCTCCA
10051 CCCCAGTAT TCCCTCTACC CCTCAGCTTT GCCTGTAAGT TTCTTTTAA
10101 AAAAATCACA CATACATTGT TGTAAGTATG TTAAGAATAA GTATTTTGCT
10151 GACCCCAAGG TCTTTTGCTT CTTTCTAGAT CAGTGCCTG CAGGTTTAT
10201 TTTACAGAGC TTAATAGAAT CAGAAATCTC TTTAAACTC CAGTCTCATA
10251 TCCAGTTATC ACTCACCATC TCTGTGTTG CAGCAATAGC CAGGCCTGGC
10301 CCAGAGGGAC TTGATCTCCA CTTTGGTTT TTAGACTTTT CTGTGGCTTT
10351 TACCACCTGC TGTGTATCCT TGACCTCATA CTGCTGGACT CCTTTGGATG
10401 GATACCAGCA GGATGGTTCA GGCTCCAGTG GGCACTTTT AAAATTCTCT
10451 CCTTCTGTTT AGATAGACAG AGCTCAGGCA GATCACCAG TCTGTTGCC
10501 GTGTAACCAG GAAAGAGATG CTAGTTTCT TTTAGGCACT CCCATTTGTT
10551 TCTGTTGGAA CCTTCTCAC TTAGTTGATG GAAGGGAAGC AAAAGACCCA
10601 GAACTCCATC TCAAATTAAT GACTTAACAA TTCTTGAATT TTCTCTTATC
10651 TCCTAGTTAA CTCTTTTCT TATCTCCAGG AGGTCAGTTT TAATTATTGT
10701 TGTTTATTCA TTTATTTTCA TGGAGACAGG GTCTTACTGT GTTGTCCAGG
10751 CTGGCCTTGA ACCCCTGGCC TCAAGGAATC TTTTCACCTC AGCCTCCCAA
10801 AGTGCTGGAA TTACAGGCAT GAGTCACCAC ACCCAGCCTG ATATTTTCA
10851 GTTGATGTAT CATAGTTGTG CCTAAGCATA ATTTTAAAT TTAATTTT
10901 TATTTTGGG GACAGGGTCT CCCTCTGTCG CCCAGGCTGG AGTGCAGTGA
10951 TGCGATCTCA GCTCACTGCA ACCTCCACCT TCTGGGTAA AGCGATTCTC
11001 TGGCTCAGC CTCCCGAGTA GCTGGGACTA CAGGTACCCA CCATCACACC
11051 CGGCTAAAT TTTTGTGTGT ATTTTAGTAG AGACGAGATT TCGCTGTGTT
11101 GCCCAGGGCG GTCTCGAACT CCTGAGCTCA GGCAATCCGC CTGCCTCGGC
11151 CTCCCAAAGT GCTGGGATTA CAGGCATGAA CCACCACGCC CGGCCAAGCG
11201 TAATATTTT AAGGGTCATC AATGTTGTGT CATGAATCAA TCAGTGTTC
11251 GTTCTTTTT ATGGTTGAAT AATATTCCAT GGTATGGATT GTGCACATTT
11301 TGTTTATCCA TTCATTAGTT GATAGACATT TTGGATTCC ACTTTTTTTT
11351 TTTTTTTTT GCTATTATAA ATAGTGATAC TATGTACAA TTTTGTGTG
11401 GAAATATGTC CTCATATCTC TTGGTTATAT ACCAAAGAGT GGAAGTGCTG
11451 GGTCATATGG TAACTACGTG TTAACATTT TGAGAAACTG CTAACCTGTT
11501 TTCCAAAGTT GCTGTACCGT CGTACATTCC TGCCAGCAAT ATATGAGGAT
11551 GCCAGTTCCT TCACATGTTT ACTACACTTA TCCACCTTTT TTATAATAAC
11601 TAATGGTGGG TGTGAGATGG TATCTCATTG TAGTTTTGAT TTGTATTTCT
11651 CTGATGGCTA AATGGCTAAT GATGTTTGAA CTTTTGTTT GAGACAGAAT
11701 CTCACCTCTG CCAGATTCAA GCGATTCTCC TGCCCTCAGCC TCCCTAGCAG
11751 CTGGGATTAC AGGCACATGC CACCACACCC AGCTAATTTT TTGTATTTCT
11801 AGTAGAGACA GGGCATTACC ATGTTGTTCA GGCTGGTGTG GAACTCCTGA
11851 CCTCAAAGGA TCCGCTCCC TGGGCTCCC AAAGTGCTGG ATTACAGGCT
11901 AGAGCCACCA TGCCAGGCTT TATGTTTGAA CATCTTTAT GTGCTTATTG
11951 GACATTTGTG TATCTTCTTT GGAGAAATGT CTGTTCAAAG TCTTTGTCCA
12001 TTTTAAATTG GATTGTCTTT TTGTCTTTG ATGTGTAAGA GTTCTTTATG
12051 TGTTTTGGAT ACAAGTTTGT TAGATATATG ATTTGCAAA CTTTCTCCA
12101 ATTTTTGTGG ACTTTTGCTT TCTTTTTTG TTTTGTTTT GTTGTGTTG
12151 TTGTTGTTGT TGTTTTGGTC GGGGGACAGT CTGCTCTGA CCACCCAGGC
12201 TGGAATGGAG TGGCGGATC TTGGTTCACT GCAACCTCTG CCTCCTGAGT
12251 TCAAGCTATC CTGCTCAGC CTCCCGAGTA GCTGGGACCC AGGTGTGTGC
12301 CACCACTCCC AGCTAATTTT TTATTTTATG TAGAGACCGG GTTTCACCAT
12351 GTTGGCCAGG CTGGTCTTGA ACTACTGACC TCAGGTGATC TGCTGCCTC
12401 AGCCTCCCAA AGTGCTGGGA TTACAGTCAT GAGCCACTAC ACCCTGATTC
12451 TTTTGTCTGG CTTTCTTTCT TTTTTTTCT TTTTTTTTT GAGACGGAGT
12501 CTCGCTCTGT TGCCAGGCTG GAGTGCAGTG GCATGATCTC GGCTCACTGC
12551 AACCTCTGCC TCCCGGTTT AAGCCATTCT CCTGCCTCAG CCTCCGAGT

FIGURE 3, page 4 of 17

12601 AGCTAGGACT ATAGGCACAT GCCACCATGC GCAGCTAATT TTTGTATTTT
12651 TAGTAGAGAC GGGGTTCAC CATGTTGGCC AGGATAGTCT CGATCTCTTG
12701 ACCTCGTGAT CCGTCCGCTT GGGCTCCCA AAGTGCTAGG ATTACAGACG
12751 TGAGCCACCA CACTCAGCCT CTTTTTGCTT TCTTGATGGT GTCTTTTGAA
12801 ACAAAGTTT TTACTTTTGA TAAAGTCCAA TTTGTCTATT TTGTTTGT
12851 GTTTTGTGTA AGAAGCTTTG CCTAACCCAA AGTCACGAGA ATTTTCTCTT
12901 AGGTTTCTT CTAAGAGTTT TATAGTTTGA GCTGTTTCTG TGATCCATTT
12951 TGAGTGAATT TTTGTGAATG GTATGAGGGA GTGATCCAAC TTCATTCTTT
13001 TGTGTGTGGA TATCAAGTTG TCCAGCACT ATTTGTTTAA ACCACTGTTC
13051 TTTTCCCCCA TTGAATTATC TTGGCATCAT TGTCAGAGAT AAATTGACCG
13101 TAAATGTGAG GGTTTTATTT CTGAACCTC AAGTCCATT CATTTGGTCTA
13151 CATGTCCCTA TGCCAGTAAT ACATACTCTT GGTTACTGTA GCTTTTGTAGT
13201 ACGTTTGGAA ATGTTTTTAA AATTTGTTTT TCATCTAAAT TTTAGGATTA
13251 ATTTGTCAAT TTCTGCACAA AAGGCACCTG GGTTCTATA GGGGTTATGC
13301 AGAATCTGTA GATCAACTGG GGGAGTATTA CAGGCATGAG CCACCGTGCC
13351 TGCTGACTG AGTTTTTCAT AGATGTACTC TATCAGGTTT AGGAAGTTCC
13401 CTTTTATTCC TAGGTGTGTT AGTCTATTTT ATATTACTTT TTTAGAGACA
13451 GTCTTGCTCT GTCCCTCAGG CTGTAGCACA GTGGCTCAAT CATAGCTCAC
13501 TGCAGCCTTG AACTCCTAGG TTCAAGAGAT CCGCCTGCCT CAGCCTTCTT
13551 AGTAGCTGGG ATTACATGCA TGCACCACCA TACTGGGCTA ATTTTAA
13601 ATTTTATA GAGACAGGGT CTTATTACTA TGTGCCCAG ACTGGCATTG
13651 AGTCTTTTGA TCATTAATGA GCACTGAATT TTGTCAAGTG CCTTTATAAT
13701 ACCTATTGTG ATGATCATAG GGTTTTGTTC TTTAGTCTAC CGATACGCTA
13751 TATTGCATTA AGTGATTTTT TTGAATGTTA AACCAACCTT GCATTTTTTT
13801 GGTGTATAAG TCTTATTTGA TCAATGTGTA TTATCCTTTT ATATGGTGCT
13851 GGATTTAGTT TGCTACTATT TTGTTGAGGA TTTTGTGTC TATATTCTA
13901 AGAGATATTG GTCTGTAGTT TCTTGTGATG TCTTGTCTG GTTTTAGAAT
13951 CAGGGTAATG CTGGCCTCAT AGAATGAATT GGGAAAGTGT GTCTTTTCTA
14001 TGTGATGGGA GAGTTGTGTA ATCATTGCTA TTAATTTTTC TGTAAATGTT
14051 TGGTAGAATT CACAAATAAA GGCATCTGAG CCTGGGCTCT TCTTGTGGG
14101 AAGTTTTTGG CTTTTTTTTT CTTTAAAAAT TTTTATTGTG GCTGGGCATG
14151 GTGGCTCACG CCTGTAATCC CAGCACTTTG GGGGGCCAAG ACGGCTGAAT
14201 CAGCTGAGN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNGAG
14501 TGCAATGATT GCAGTCTTGG CTCACTGCAA CCTCTGCCAC CTGGGCTCAA
14551 GAGATTCTCC TGCCCTCAGC TCCTGAGTAG CTGGGATTAC AGGCATGCGC
14601 CACCATGCCC AGTTAATTTT TGTATTTTGA GTAGAGACGG GGATTCTCCA
14651 TGTGACCAG GCTGGTCTCT AACTCCTCAC CTCAGTGAT CCGCCCGCCT
14701 CGGCCTCCCA AAGTGCTGGG ATTACAGGCA TGAGCCACCA CGCCCGGCCT
14751 TAAAAATTTT TTTAATGTAC AGTTGAGTAG TATTTAATAC ATTCACATTG
14801 TTGTGTACCC AGTTTCCAGA ACTCTTCATC CTACAGAAGT GAAACTCCAT
14851 ACCCATTAAG TGAGTCCCCA TTCTCTTTCC CCCAGCTCAT GGCAACAGC
14901 ATTCTATTTT CAGTCTCTAT GAATTTGATT AGTTTAGATA CTTTACTG
14951 TAAGTGGAAT CATATGGTAT TTGTCTTTTA GTGACTGCCT AATTAAAAA
15001 AAATTTTTTT GAGACGGAGT CCTGCTCTGT CGCCAGGCT GGAGTGCAGT
15051 GGCACCATCT CTGCTCACTG CAACCTCCAC CTCCAGGTT CAAGTGATTG
15101 TCCTGCCTCG GCCTCCACG TAGCTGGGAT TACAGGTGCT CGCCACAACA
15151 CCCGGCTAAT TTTGTATTTT TTAGGTAGAG ACTGGGTTTC ACCATGTTGG
15201 CCAGGCTGGT CTCGAACCTC TGACCTCAA TTATCCACCT GCCTTGGCCT
15251 CCCAAAGTGC TGGGATTACA GGCCTGAGCC ACTGTGCCCA GCCTCCATGT
15301 TGTTTTTTCAC AACACCTGTA TCATTTACAT TTCCACCAAC AGTACACAAG
15351 AATTTAGTT TCTCCACATC CTTGCTAGCA GTTGTATTA TCTGTTTTTT
15401 TTTAATGGTT TCTTTTCTCC TTTTCTTTT TTTTCTTTT TGAGACGGTC
15451 TTATTCTTGC TCATGCTGGA GTGCAGTGGT GCAATGTGAT AGCTCACTGC
15501 AGCCTCAACC TCTGGGCTCA AGCAGTCTG CCACCTCAGC CTCCACATAG
15551 GTGGGACTGC AGGTGTGCAC CACCACTTGT GGCTAATTTA AAAAATTTTT
15601 TCGTAGAGAC AGAGTCTCAC AGTGTACCC AGGCTGGTCT TGAATTCTG
15651 AGCACAAGTG ATCCTCCAC CTCAGCCTCC CAAAATAATG AGATTAGAGA
15701 CATGAGCCAA CATGCCAAC CAGTTTTGTT TGTTTGTGTT GTTTTGTGTT

FIGURE 3, page 5 of 17

```

15751 TTTGAGACAG AGTCTCACTC TATTGCCAG GCTGGAGTGC AACGGCATGA
15801 TCTTAGCTCA CTGCAACCTC CGCCTCCCAG GTTCAAGTGA TTCTCATGCC
15851 CCAGCCTCCT GAGAAGCTGG GATTACAGGT GTACCACCAC ACCCAGTTAT
15901 TTTTGTATTT TTAGTAGACA TGGGGTTTGG CCATGTTGGC CAGGCTGGTC
15951 CCGAACTCCT GACCTCAAGT GATCTGCTCC CCTCAGCCTC CCAAGGTGCT
16001 AGAATTAAGT TTTTCTTTCT TTCTTTCTTT CTTTTTTTTT TTTTTTTTGA
16051 GACAGAGTCT CACTCTGTCA CCCAGGCAGG AGTGCAATGG CACGGTCTTG
16101 GCTCATTGTA ACCTCTGCCT CCCAGATTCA AGTAGTGATT CTCCTGTCTC
16151 AGCCTCCCAA GTAGCTGGGA TTACAGGCAT GCACCACCAC GCCCAGCTAA
16201 TTTTTTGTAT TTTTAGTAGA AACGGGGTTT CACCATGTTG GTCAGGCTGA
16251 TCTCAAATC CTGACCCCAA GTGATCCACC CGCCTTGGCC TCCCAAAGTG
16301 TTGGGATTAC AGGCGTGAGC CACTGTGCCT GGTTTTATTT TTATTATTAT
16351 TATTTTTAAT AGTTCCTATT CTAATGGGTA TGAGGTAGTG AGGTGGGTGG
16401 TTTGGGTGTT TTTATGAATG TTTAATTGGA AATGGGTGGC CATTGTGTGC
16451 AGGAAAAACC TCCTAAATTG TGTCAAATC CTGGAATATG AAATATCATT
16501 CCAGTTGCAA GAATATCTTT TTTTTTTTTT TTTTTTTTTT AGACAGAGTC
16551 TCACTCTGTC ACCAGGCGGG AGTGCAAGTG CACGATCTCG GCTCACTGCA
16601 ACCTCCGCCT CCTGGTCCGC CTCCCGGGTT CAAGTGATTC TTCTGCCTTA
16651 GCCTCCCAAG TAGCTGGGAC TACAGGCGCG TGCCACCACT CCTGGCTAAT
16701 TTTTTTGTAT TTTTAGTAGA GATGGGGTTT CACCATGTTG GCCAGGATGG
16751 TCTCAATCTC TTGACCTAGA GATCCGCCTG CCTCGGCCTC CCAAAGTGGT
16801 GGGATTACAG GGGGGTCACC GTGCCCAGCC ACAAGAAGAT CTTGAGCATG
16851 TGAATGATCA GAAATGATTT AGCCTATGTA GGCACTAGGC CAGGTAGTGA
16901 AATTCAGGGA AAATAATTCA GATGCTTCTG AGCTATCACT TATGAACTAA
16951 GAAACAGCTT AAAGCCATTA TAGTGTGTTT CCTGAAGATG AAAGCATATG
17001 GTAAGATGAA ATAGTGATTA TTTTTTAAAT ATTACTACTC CAGAAAGGAA
17051 AAGTTTACTA ATTTTATTA CTAAAGTTA CTGTTGGTGG GTGCGGTGGC
17101 TCACACCTGT AATCCCAGCA CTTTGGGAGG CCGAGGCAGG CGGATCACCT
17151 GAGGTCAGGA GTTCGAGACC AGCCTGACCA ATATGGTGAA ACCCATCTC
17201 TGCTAAAAAT AAAAAATTAG GCCGGGCGCG GTGGCTCATG CCTGTAATCC
17251 CAGCACTTTG GGAGGCCGAG GCAGGTGGAT CACGAGGTCA GGAGATTGAG
17301 ACCATCTTGG CTAACACGGT GAAACCCCGT CTCTGCTAAA AATACAAAAA
17351 TCAGCCGGGC GTCTTGGCAG GCACCTGTAG TCCCAGGTAC TCAGGAGTTT
17401 TGAGACGGGA GAATGGCGTG AACCCGGAAG GCGGAGCTTG AAGTGAGCCG
17451 AGATTGCGCC GCTTCAGTCC AGCCTGGACG ACAGAGTGAG ACTCTGTCTC
17501 TAAAAAATA AATAAATAAA AATAAAAAAT TAGCTGGGTG TGGTGGCAGC
17551 CACCTGTACT CCCAGCTACT CGGGAGGCTG AGGCAGGAGA ATTGCTTGAA
17601 CCCGGGAGAT GGAGGTTGCG GGGAGCCAAG ATTGCGCCAC TGCACTCCAG
17651 CTTGGCGACA GAGTGAGACT CTTTCTCAGA AAAAAATATG ATAATTAAAA
17701 GTTGAGACGT TCTTCGCGGA GAGTGGTCGG GGTTTCCTGC TTCAACAGTG
17751 CTTGGACGGA ACCCGGCGCT CGTCTGCAC CCCGGCCGGC CGCCCATAGC
17801 CAGCCCTCCG TCACCTCTTC ACCATGCCCT CGGACTGCCC CAAGGCCCCC
17851 GCCGCAGCTC CAGCGCCGCG TAGCCACCAC TGCCGCTGCC GCCGCTCTC
17901 CTTAGTCGCC GGCATGACGA CCGCGTCTAC CTCGCAGGTG CGCCAGAACT
17951 ACCACCAAGA CTCAGAGGCC GCCATCAACC GCCAGATCAA CTTGGAGCTC
18001 TACGCCTCCT CCATTACCT GTGCGTGGCT TACTACTTTG ACAGCGATGA
18051 TGTGGCTTTG AAGAACTTTG CCAAATACTT TCTTCACCAA TCTCATGAGG
18101 AGAGGGAACA TGCTGAGAAA TTGATGAAGC TGTAAGACCA ACGAGGTGGC
18151 CGAATCTTCC TTCAGGATAT CAAGAAACCA GACTGTGCGG GGAGAATGCG
18201 ATGGGAGAGC GGGCTGAATG CGATGGATTA CATTGGAATA AAATTGTGCA
18251 TTTTGCATTA CATTGGGAAA AAAATGTGAA TCAGTCACTA CTGGAAGTGC
18301 ACAAACTGGC CACTGACAAA AATGACCCCC ATTTGTGTGA CTTTATTGAG
18351 ACACATTACC TGAATGAACA AGTGAAGGCC ATCAAAGAAT TGGGTGACCA
18401 CGTGACCAAC ATGCACGAGA TGGGAGCGCC CGAATCTGGC GTGGCAGAA
18451 ACCTCTTTGA CAAGCACACC CTGGGAGACA GTGATAATGA AAGCTAAGCC
18501 TCAGGCTAAT TTCCCATAG CCATAGGGTG ACTTACCTTG TCACCAAGGC
18551 AGCGCATGTA TGTGGGGTT TCCTTTACCT TTTCTATAAG TTGTCCAAG
18601 ACACCCACTT AAGTTCTTTG ATTTGTACCA TTCTTCAAA TAAATAAATT
18651 TGGTACCCTC CCCCCCCCCA AAAAAAAAT GTACTGTGGG CTGGCGTAGT
18701 GGCTCATGCC TAAATCTCAG CACTTTGGGA GGCTGAGGCG GGAGGATCAC
18751 CTGAGGTGCG GAGTTTGAGA CCAGCCTGGG CAACATGGTG AAACCCCGTG
18801 TCTACTAAAA ATATAAAAC TAGCCAGTCA TGGTGGCACA CACCTGTAAT
18851 CCCGCCACT TGGGAGGCTG AGGCATGAGA ATCACTTGAA CCTGGGCTGC

```

FIGURE 3, page 6 of 17

18901 GGAGGTTGTA GTAAGCTGAG ATCATGCCAC TGTACTCCAG CCTGGGTGAC
18951 AGGGAGACAT TCTCTCTCTC AAAAAAAAAA AAAAAACAAA AAAAAACAAA
19001 CAAACCAACA AAACAAAGTA ATCCAGGAAC AACAAACATGA TGAAGGACTG
19051 CATGCAGGAC TCAGTGATGG ATGGTGGGAG ACAGCCAGGA AGTTAAGCAT
19101 GACTCTGGTA TTAAGTGTG TCTGGGAGAG TTAAGATTCC ATTTACAGAA
19151 ATAAGACCTG TAGGGGAAGC TCTTGATTTT TTTTTTTTTG CAGACTGCTG
19201 ATTTCCCTGAT TACATGTGTT AAGTTTGAGG TATAGAGAGA AAGAACATCC
19251 TGGCCGGGTG CAGTGGCTCA CACCCGTAAT CCCAGCACTT TGGGAGGCCA
19301 AGGTGGGCAG ATCAGGAGGT CCAGGAGATC GAGACCATCC TGGCCAACAT
19351 GGTGAAACCC CGTCTCTACT AAAAATACAA AAATTAGCTG GCGGTGGCGG
19401 CGCGTGCCTG TTATCCCAGC TACTCAGGAG GCTGAGGCAG GAGAATTGCT
19451 TGAACCCGAG AGGCAGAGGT TGTGATGAGC CGAGATCGCG CCACTGCACT
19501 CCACCTTGGC AACAGAGCTA GACTCTGTCT CAAAAAATAA AAAAAGAAAG
19551 AAAAAAAGAA CATCCTGTAG AAACAGGCAG TCAGAGGTAT AGAACTACAC
19601 AGAATCCAAG AGATCTTTCA AGAAAAGTGA CATGCAGCAA GAGAACTAT
19651 CAAGGGGGTA AACACCTAT AGAATGGGAG AAAATATTCA CAAAGTATAC
19701 ATCCAACAAA AGTCTAATAT CCAAAATCTA TAAGGAACCT AACAAAGCAA
19751 AAGCAAATAA CCCCCTTAAA AAGTGGGCAA AGGACATGAA CAGATACTTC
19801 TCAAAAGACG TACATGTGGC CCACAAACAT GAAAAACGC CCATTTCTAA
19851 TCATCAGACA AATGCAAAAT AGAACCACAA AGAGATACCA TCTCACACCA
19901 GTCAGAACAG CTTTTGTAA AATGTCAAAA AATGAGAAAC GTTGGTGAGG
19951 CTGCAGAGGA AAGCAGACAC TTGTACTACT TTGGTGAAGG TGTAAATTAA
20001 TTTAGCGTAG GCACAGTCAG TTTGGAGATT TCTCAGAGAA CTAAGAGTGG
20051 AACTACCATT AGACCCAGCA ATCTGATGGC TGGATATACG GCCAAAGGAA
20101 AATAAATCAT TCTGCCAAAA GAACATATGT ACCTGTATGT TCATTGTGGC
20151 ACCATTACCA ATAGCAAAAA CATTGAATCA ACTCATGTGC CCATCAGTGG
20201 CGGACTAGAA AAGAAAAGAA AATATGGTAC ATAGCCATCA TGGAACTACTA
20251 TGCACCCATT AAAAATAATG AAATAATGTC TTTGCAACAA CATGAATGTA
20301 GCTGGAGGGC ATTATCCTAA GCAAACCTAC ACAGAAACAG AAAACCAAAT
20351 ACTGCGTGT CTACACGAGT GAGAGTGGGA GCTGAACATC AAGTACACAT
20401 GGATGTAAGG ATGGCAACAA TAGACATGGG TCTACTAGAG GTGGTGGTGC
20451 GGCAGGGTGG GGGTGGGGGT TGTGTGGCAG AGGAACAGCT GAAAAACTAC
20501 CTATTGTGATA CTATACCCAG CACCTGGGAA ACGGGTTCAG TCATACCCCA
20551 AACCTCAGCA TCACACAGTA TACCTTTCTA ACAAACTTAC ACATGTATTTC
20601 TGTGATTCTA AAATAAACAT TGAAAAATAA AAAAAAACT GACATGGTTT
20651 GTAATGTTTA ATCTGACATA ATGGCTAGGG GAAATGAAGT CTGCAGAATG
20701 GCTGTTTACG GATGTTGTTG TTGTTGTTGA GATGAGGTCT CACTATGTTG
20751 CCCAGGCTTG AACTCCTGGC CTCAAGCAGT CCTCCTGCCT TGACCTCCCA
20801 AAATGTTGAG ATTACATGCA TGAGCCATTG CCAAAACGGC TATTGGATT
20851 GCTGTTAAGG TTATTACATT CTCTGTGTAG TAAGACCTTG AAGGAGAAGG
20901 ATTTGAGATC AGGAGTTTAA GAAAAATGT TAATCTAGGA AGAGAGGATA
20951 ATTTCTGTGT TTGGCCAGTT GCAATGGCTC ACGCCTGTAA TCCCAGCGTT
21001 TTGGGAGGCC GAGCTGGGCA GATCACTTGA GCTCAGGAGT TTGAGAAGAG
21051 CCTGGGCAAC ATGGTGAAGC CCCGTCTCTA CTGAAAATAC AAAAATTGGC
21101 TGGGCGTGGT GGCAGGTGCC CACTGTAATT TCAGCCACTC AGGAGACTGA
21151 GGCAGGAGAA TTGCTTGAAC CTGGAAAGTG GAGGATATGG TGAGCCTAGA
21201 TCGCGCCACT CTACTCCAGC CTGGGAGACT CCATCTCAAA AAAAAAGAA
21251 AACAAAGATG TGAAATGAAG TAATTACCAC AGTCAATGTG ATCCTATAAC
21301 TTTGTTTCT TTTAGAGATG GGGTCTCCCT CTGTACCCCA GGCTGGAGTG
21351 CAGTGGTGCA TCATAGCTTT CTGCAGCCTC CACCTCCTGG GCTCAGGTGG
21401 TCCTCTTGCC TCAGTGTTCC GAGTAGTTAG GACTGACTGC AGGTGCATGC
21451 TGCTATGCCT GGCTAACTTT AAAATTTTTT TGTAGAGGCG GGGTCTTGCT
21501 ATGTTGCCTA GGCTGGTCTC CAACTCCTGA TCTCAATCAG TCCTCCTGCC
21551 TACCTTCCCA GAGCGCTGGG ATTACAGGTG TGAGCCATCG CACCTAGCCA
21601 ATCTCATAAC ATTTTATGAC TAGCAAACCT AGTAGTTCTG ATTCAGGCAT
21651 AAATCAGTTG GTGGGGTTAT ACAAGGTTGG GTGAGTTTTT CTAGATTTCT
21701 AAGAGACCAT GTTGAAATAC TTGGCTCTGT CTCAGTAAGG GATAGAGAGA
21751 AGCAAAGGTG TGGGTAAAGG TTATGAGCAG ACATGTAAAG GGGGCAAATT
21801 AAAGTGTFTA GGGGAAGGTGA AACAGTTCCA AATCATAATA TAGATCCCAA
21851 GCCTACCAA GAAGTGAAGG AGAGGGTAAT TGTGCGATAG TTCTCAGGAC
21901 TGAGACCTCA AGGTGTAAGA CGAATCTTTA TTGTGGGTGG TCCTCTTTGA
21951 GAAAAAGAAC AAAAAAGAAA AATGTGAAAT GGGTGCTAAA TTTAGAGCA
22001 GAATATACAT ATATGTGTAT GTGTGTGTGT ATATATATAT ATATTTTTAT

FIGURE 3, page 7 of 17

```

22051 GTAGCAAATT CCAGTTTACA AAGGGCTTTT ATGTGTTTGT CATCATTACA
22101 ACAGTTCGTGAT GATGATGTGG ATGTGGTGGA TATTGTAATT CCACATCCCA
22151 GAGGATAAAA TTGAGGCACA CCATAAGGTA GCTGACCAGA GATCATGCAG
22201 TATATGGTAG AATGGCAACT TGAGGCCAGT TCTTAGGTGT GTTGTGTGTC
22251 CTTATTCAGA ATGGAACAAT GTGGTTTATT GTACAAAAAT TAAAAATGA
22301 ATGTCGAAAA GTAGAAATTA TTACCCAAAT CTCACCCTCT GTGGTTGTTT
22351 GCTATGTGAT CTCCAGACT CACATATACA TGTAGATAAT TTTTTTTTTT
22401 TTTTGTGAGA CGGAGTCTTG CTCTGTCGCC CTGGCNNNNN NNNNNNNNNN
22451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
25001 NNNCAAAAT TAGCTGGGCG TAGTGGTGGG CGCCTGTGGT CCCAGCTACT
25051 TGGGCAGCTG AGGCAGGAGA ATCGCTTGAA CCTGGGAGGT GGAGGTTGCA
25101 GTGAGCCAAG ACCATGCCAC TGCCTCCAG CCTGGGCAAC AGAGCAAGAC
25151 TCTTGTCCTTA AAAAACAAAA AAGTGTACCC AGTTGAGCTG ATTCCTTATC

```

FIGURE 3, page 8 of 17

```

25201 TTTTTCAC TGGAGAACTA AGTATACAGG TGAGAAAAGA CGAGATATTT
25251 ATACCCGAGA GAATTGATGG TGAAATCCAT TTTTGGAT CAGAACTTCC
25301 CCAAACAGTG TCCTTCAAAT AGGGTTCAGG GGTGCTAAGA TATTATCCC
25351 CTCAACCCCTT GGGGTTCACCT CCAGTATGGC ATATAAATAT TGTATCACTT
25401 TCTATGTGTG GGGAGCAGTG CTCCAGGTGA CCTTCCTTCC TTTCTTCTA
25451 GGGGAGGAAG TTGCCTTGTA CTGTGCCAAA TATCTTCCTG ATATCATCAA
25501 AGATCAGAAAG GCCTACAAGG AAGGCAAGCT ACAGAAGGTC TGTCTGCTTA
25551 CACCGCCCAT TCCTCACTTG TGTAGGCTTT TCCCTTGTTT TCTAGCCCTT
25601 GGGCTTTTCC TTTCTTTTGT TCCTCTAGCT GCTGCTGCTT ATTTACTCTT
25651 GAAGAATTCT GTTCTTAAAA CGAGCTTATT GGCCGCCTTT TAGACTTGCC
25701 TTATTATTCC TAGGCCTCTG AGCTGTTTTT ATCTGTGAGT GTCTCTTAGT
25751 GTGGTGGCTC ACACCTCTTA TTTGTATTCC ATCCTTGTCG TCAGGATTGT
25801 ATATAGGGAG TTCATTTTGT ACTAGTCTTA GACTATTTTG CTTATATTCA
25851 GGCTTTAGAA GATGCCTTCT TGGCTATTGA CGCCAAATTG ACCACTGAAG
25901 AAGTCATTAA AGAGCTGGCA CAGATTGCAG GGCGACCCAC TGAGGATGAA
25951 GATGAAAAAG AAAAAGTAGC TGATGAAGAT GATGGTGAGT GTGGCATCCC
26001 TTGTTTGAGG GAAATCAGC ATTTTAAGAA ATATTCTTTA ATATTACTTA
26051 TCAATTCTAA GATAGGATGG CTTTCTAGGG ACCTGGGGAG TCCTTATGTT
26101 AAAGAAACCT ATGATGTTCT CCTGCATTGT ATGTGGTTAT GAAAAGGAGG
26151 GAGAGAATTA TCTTCGTTGA GTGGCATCTG AGCTGTAAAG ATTGTATATA
26201 CATTATCTTT TGTCATTGTG ATGGGGTCTT CCTGGTTCCT GCTAGTATTT
26251 ATGTGCTTTT TTTCCCCTC AAGACTGGAG CAGTTATTAG CCCCAATAGC
26301 CAATCATTAA GCCTAAATCC TAATTCACAG TAGCATTGTG GGCTTCCTGG
26351 ATCCTCAGCC AGAATAGGGT TTTTACAAC TAACAATAAA AAATGAGACG
26401 TCAGAGGGGA AGTATAGTAA CTAGTGTGTT TTTGATTAAG AAGGGGATGA
26451 AACACAAAAA CCAAAAGAAG TCTGTGGAGG AGGAGGAGCT AGGGCATGTT
26501 CTTCTGAGAC TTGAGCGAGA GGAACCTTGG GAGTGGGAGG TTGTGGGGAA
26551 GTTAGAGGCT GCAAGGGCTG TTGAGGTAGT GAGAGGGACG GATCCCATGA
26601 GGAGTCTGGC ATGGGGGGCTC TGATTAGGCC TCTTCCCTGC AGTGGACAAT
26651 GAGGAGGCTG CACTGCTGCA TGAAGAGGCT ACCATGACTA TTGAAGAGCT
26701 GCTGACACGC TACGGGCAGA ACTGTCACAA GGGCCCTCCC CACAGCAAAT
26751 CTGGAGGTGG GACAGGCAGG GAACCAGGCT CCCAGGGCCT CAATGGGGAG
26801 GCAGGACCTG AGGACTCAAC TAGGGAAACT CCTTCACAAG AAAATGGCCC
26851 CACAGCCAAG GCCTACACAG GCTTTCTCTC CAACTCGGAA CGTGGGACTG
26901 AGGCAGGCCA AGTTGGTGAG CCTGGCATTG CCACTGGTGA GGCTGGGCCT
26951 TCCTGCTCTT CAGCCTCTGA CAAGCTGCCT CGAGTTGCTA AGTCCAAGTT
27001 CTTTGAGGAC AGTGAGGATG AGTCAGATGA GGCGGAGGAA GAAGAGGAAG
27051 ACAGTGAGGT AAGGGCCTGT GAGGGCAGGC AGATGCTGAA GTTGACAGA
27101 GGTCTGTGTT GGTGCGGTC TGTAAGTTTC AACTCTCTTT CCTTCTCCTA
27151 TTTTGACATC ATCCCCAAG ACCCACTGTA TTCTAAGCTT TAGTCTTGAA
27201 TTCATTGAGC TCCATCATCA CAGGTACCAT TTGCCTTTTT ACCTCTTCCT
27251 TTGTTGGTAC TATAACAAGC AGATCTAGTT CTGGCTTTTC AGAGTCTGTC
27301 TCCTAGAGAG AGAACAAGGA GATAGTTGTT ACCTTGGCTA GTTGACTGTT
27351 TTCTTCTCTG GAAAATTTAT TTTCTGGCCA CAGTGCCTGA AAGATATTTT
27401 TGGCTGGCAG CCCTTGCCCTT GTCCTGGGCT TTTTGCTAGT GACTGCTAAG
27451 CCCAGTTCAG GATGTCAGTT GTACTCATGC TAGCCCTTTC CATCCCCCA
27501 ATTTTCATGA CCATATACTT GTATCTTTCA GTGTTTTGAG GACCTGTGTT
27551 CAGTCAGGAC CTCTTGATTG TGAGTATGAG CTGTGGGGAG GGAGGGGATC
27601 ATCCCAGTCT CAGCAGTCTG GGATCCTCCC CCTGGCAGGA ATGCAGCGAG
27651 GAAGAGGATG GCTACAGCAG TGAGGAGGCA GAGAATGAGG AAGATGAGGA
27701 TGACACCGAG GAGGCTGAAG AGGACGATGA AGAAGAAGAA GAAGAGATGA
27751 TGGTGCCAGG GATGGAAGGC AAAGAGGAGG TGTGTGGGGA AGGGGAGCAA
27801 TGAGTCTTGA AAAGCCACAA GGCAGGTGTG AATCCCTTAA TTTTGATTTT
27851 GAGACAGGGG ATCCCCCTGA TACTTTAGGA TGGAAGTAAT AGTCATGGGG
27901 ATTTATTCTG CAAGGGGAAT GAGATGGTAA GCCTTTGGGG TTGAATTATC
27951 TAAAAACAAG GGAGAGGGAG TGTGCTGCTG TCTCTAGAAA GATGAAATGT
28001 GTGCTTCTCC TGTTTGTTAA AGCTCTTTTG GGGGTCCAG TGAAAGCAAG
28051 CATAGGTGAA CGATCAGGAG CACATCAGTG AGGAACGCAT GTTCAGAAGC
28101 CCCCATGATG CTCCTTTTCT TCCTCTTAAG CCTGGCTCTG ACAGTGGTAC
28151 AACAGCGGTG GTGCCCTGA TACGAGGGAA GCAGTTGATT GTAGCCAACG
28201 CAGGAGACTC TCGCTGTGTG GTATCTGAGG CTGGCAAAGC TTTAGACATG
28251 TCCTATGATC ACAAACCAGA GGATGAAGTA GAAGTAGCAC GCATCAAGAA
28301 TGCTGGTGCG AAGGTCACCA TGATGGGCG AGTCAACGGG GGCCTCAACC

```

FIGURE 3, page 9 of 17

```

28351 TCTCCAGAGC CATGGGTAAG GGCCAAGAAA CTGGGAAAGA NNNNNNNNNN
28401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNATGGTGAA ACCCCATCTC
28751 TACTAAAAAA AAAAAAAAT ACAAAAATTA GCTGGGCATG GTAGTGCAAG
28801 CCTGTAATCC CAGCTACTCA GGAGGCTGAG GCAGGAGAAC CGCTTGAATC
28851 CGGGAGGCGG AGGTTGTAGT GAGCCGAGAT CGTGCCATTA CACTCCAGCC
28901 TGGGCTACAA GAGTGAACT CCGTCTCAA AAAAAAAAAA CAAAAAGAC
28951 TTAAATAAA AAGACCAAGT AGTGACTTTC TTAAGGTCA GCAGTCTGGT
29001 GGCAGGGTTG AACTAGAAA AACTAGGACT TAGGACTCAG TTCCCCATTCT
29051 CACTAGATTA TGGAACTTTG TAAAGAAGGG AAATGAATGG CAAGGTTTGA
29101 CCTGCCACAA ACACAAGTCT GTGGGAAGTA TCCAACTGC TCATCAACCA
29151 TTCCTTTACT CCAGGGGACC ACTTCTATAA GAGAAACAAG AACCTGCCAC
29201 CTGAGGAACA GATGATTTCA GCCCTTCTCTG ACATCAAGGT GCTGACTCTC
29251 ACTGACGACC ATGAATTCAT GGTCAATGCC TGTGATGGCA TCTGGTGAGC
29301 ACTGGCAGAA TGCCCTAAAT TCCCCTTTCT GCAGCATGTC TTCTCTTATA
29351 GGA CTCAGG CACCTCTAGG ATTAGAGCCT AGGCAGACCT AGGCCTCTTG
29401 GTGGGTGAAG AGCACCCAGA CTAAGGCAGA GCTGAGAATT TCTGTAGTTA
29451 TTTACACTGG CCTGGGCCAC CACCTCTGTC CATACTCCTC TACGCTGCCT
29501 TAGTGAGACT GGAAGATTCT GACTGTTGTT CTTGACCCCA GGAATGTGAT
29551 GAGCAGCCAG GAAGTTGTAG ATTTTCATTCA ATCAAAGATC AGCCAGCGTG
29601 ATGAAAATGG GGAGCTTCGG TTATTGTCAT CCATTGTGGA AGAGGTGAGT
29651 ACCAGGTGG AGAAGAGAGG GTGTCTGGTC TGCACAGCCA GGGTT (SEQ ID NO:3)

```

FEATURES:

```

Start:      2134
Exon:       2134-2326
Intron:     2327-25451
Exon:       25452-25537
Intron:     25538-25851
Exon:       25852-25984
Intron:     25985-29164
Exon:       29165-29294
Intron:     29295-29541
Exon:       29542-29647
Stop:

```

CHROMOSOME MAP POSITION:

Chromosome 2

ALLELIC VARIANTS (SNPs):

| Position | Major | Minor |
|----------|-------|-------|
| 2944 | A | - |
| 3232 | G | A |
| 7236 | T | A |
| 8187 | T | A G |
| 8187 | T | - G |
| 8187 | - | A G |
| 9172 | C | A |
| 9173 | T | - |
| 9180 | T | - |
| 13351 | T | C |
| 15100 | G | C |
| 15255 | A | C |
| 18996 | A | C |
| 18996 | A | C |
| 18996 | A | C |

| | | |
|-------|---|---|
| 20532 | C | T |
| 21118 | G | T |
| 21211 | C | G |
| 21212 | T | C |
| 21238 | - | A |
| 26483 | G | A |

Context:

DNA

Position

2944

AGATGTGAATAACTTTTTAAATTTTAATATTTAAATATCTGATGTGGGAAGCCTCTTTT
GGCTAGGAGTTTGACAGTGAAAGGAACCCCGGCAGAGTCTGTTTCACATTTTGGTTGCC
TGGCCTTGGGCTCTTGCATGTAAATTTAGAGGCTGGACCCGACCTCCAGGAGTTGTAC
TCATTTGCACTCTTTTACGGGCTTTTACTAACTTCGAAAACCTGAATTATGTCAGTCC
CTAGGTTTTCTTTTATTATGCTTCGTTTTCTTCTCTCTTAGGATTCTCTAAAACCT
[A, -]
ATCAGTAATTCTCATTTGCCTGTAATTGTAGTTTATAGTTTCTTTGATTGTAAGTCATTT
CATCAATTTTTTCATCACACCTACCTACAAAGGGCTTTTCTAGAAAATTTTACTCTGG
ACAAAAGGGGAAAAGAAAATATTGGGGGAAAGTAGTAGTATTAGGTAAAACCTGATGT
GAAACTACAAAAGAGAAGAGGGAAAACCTGCGGTAGGGAGGAAAGGGAGGAAGACGGGTTA
ACCGTGGCTTTGTGAAGAGCATTCTGAAGCTAGGCAAAAGGGCCAGGAAATACTCTGT

3232

TCTCTAAAACCTTAATCAGTAATTCTCATTTGCCTGTAATTGTAGTTTATAGTTTCTTTGA
TTGTAAGTCATTTTCATCAATTTTTTCATCACACCTACCTACAAAGGGCTTTTCTAGAA
AATTTTACTCTGGACAAAAGGGGAAAAGAAAATATTGGGGGAAAGTAGTAGTATTAGG
TAAAACCTTGATGTGAAACTACAAAAGAGAAGAGGGAAAACCTGCGGTAGGGAGGAAAGGGA
GGAAGACGGGTTAACCGTGGCTTTGTGAAGAGCATTCTGAAGCTAGGCAAAAGGGCCAG
[G, A]
GAAATACTCTGTCTGGTATTGAGGGTTTCTCCACCTACCGGGTGGGCTTCAGGTAACAGC
GAAATACTGTCTCCCTTGGGAATTGTTTCAGATCCCTCGCTCCTCCTGTGGTTAGCTCTG
GAATGCCAGTATGAACTTCAATGTTTGTGTTTCCGATTCAAATTTTATATTCACTAGTGA
CCTTAATAACAATTTTACAATTAGGTATAAAATTTAGGATCCTAGTGTATCCTATAGTT
CATCTCATCTGCTTTGGCTCCCTTTTTTTTTTTTTTGTAGACAGAATTCGCTCTTGTGGC

7236

GACAGGATTTCACTATGTTGGCCAGGCTGGTCTCAAGACTCCTGACGTTGTGATCCACCT
GCCTCGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCTACCGTGCCCGGCCAGGAATT
TTTTGTGCTATAAATCATATTTTCCCTTTTATTAAGGCAGTGTCAATATCTATAGTATAA
TTTTGAGGAGGCTGGCTATTTATTTGCTGTGTAGAAGCTGGCTTATTAGTGGTCAAGGGGT
CATCTAGAATTGACTATAAAGATAGTATTGAGCAGAAAATCTTAAAATACCTGCATATT
[T, A]
GTTTCAGTCATTAAATTAATGAAAAAATATAAAAAGAAATATCACAAAGTATGCTATGGG
GTTCTACCTTAGGGCATTGAAGGTTGAAAAACATTTATTTTCTTATCTTCAGATTAGCAT
CTCATATCAATAGCCAATAGCTTAAAGCGCTTTTACTTACTAAACCAGGTGAGAAATTC
TCTCTCTCTCTCTTTTTTTTTTTTGTAGACAGAGTCTCTGTACCCAGTCTAGAGTGCAGTG
GTGTGATCTCGGCTCACTGGCCTTTCTGTTAAGTGATTCTCCGGCCTCAGCCTCCTGAG

8187

GTTGCCCAGGCTGGAGTGCAATGGTGCCATCTCGGCTCACTGCAGCTTCTGCCTCCTGGG
TTCTAGCAATTTGCCTACCCAGCCTCCCAAGTAGCTGGGATTACAGGCACTCGCCACCA
TGCCAGCTAACTTTTTGTATTTTGTAGAGAGGGGTTTTCCCATGCTGGCGAGTCT
GGTCTTGAACCTTCTGGCCTCAAGTGATCTGCCCGCCTCAGCCTCCCAAAGTGCTGGAATT
ACAGGCATGAGCCACTGTGCCCGCCTTTTTATTTTTTAAATTTATTTATGTATTTATTTT
[T, A, G]
AGACAGGATCTCACTCTTGCCCATGCTTGAGTGGAAGGAGTATGGGATTTCTTGTGCCG
GTTCTTCACTACCTCACTACCTTATCTGCCTTTACAGTGGCTCGATCATAGTTCACTG
CATAGCCTTCTGGGCTCAAGGGGTCTTCCAGCCTCAGCCTAATATAGGCACATGCCACCA
TGACTGGCTAATTTTTTTTTTTTAAAGTTTTTTTTTGTAGAGATAGGGCTTGCAGTGTG
CCCAGGCTGAGGAATTTATTTATGTTTATTTTATTTATTTATTTATTTATTTATTTATTT

8187

GTTGCCCAGGCTGGAGTGCAATGGTGCCATCTCGGCTCACTGCAGCTTCTGCCTCCTGGG
TTCTAGCAATTTGCCTACCCAGCCTCCCAAGTAGCTGGGATTACAGGCACTCGCCACCA
TGCCAGCTAACTTTTTGTATTTTGTAGAGAGGGGTTTTCCCATGCTGGCGAGTCT
GGTCTTGAACCTTCTGGCCTCAAGTGATCTGCCCGCCTCAGCCTCCCAAAGTGCTGGAATT

FIGURE 3, page 11 of 17

AGGTTGTTGAGTCTATTTTATATTACTTTTTTAGAGACAGTCTTGCTCTGTCCCTCAGGC
 TGAGCACAGTGGCTCAATCATAGCTCACTGCAGCCTTGAAGTCTAGGTTCAAGAGATC
 CGCCTGCCTCAGCCTTCTTAGTAGCTGGGATTACATGCATGCACCACCATACTGGGCTAA
 TTTTTTAAAAATTTTTTATAGAGACAGGCTTATTACTATGTTGCCAGACTGGCATTGA

15100 GTTGTGTACCCAGTTTCCAGAACTCTTCATCCTACAGAACTGAACTCCATACCCATTAA
 ATGAGTCCCCATTCTCTTTCCCCAGCTCATGGCAAACAGCATTTCTATTTTCAGTCTCTA
 TGAATTTGATTAGTTTAGATACTTCATACTGTAAGTGGAAATCATATGGTATTTGTCTTTT
 AGTGACTGCCTAATTTAAAAAAATTTTTTTGAGACGGAGTCTGCTCTGTGCGCCAGGC
 TGGAGTGCAGTGGCACCATCTCTGCTCACTGCAACCTCCACCTCCAGGTTCAAGTGATT
 [G, C]
 TCCCTGCCTCGGCCTCCACGTAGCTGGGATTACAGGTGCTCGCCACAACACCCGGCTAAT
 TTTTGTATTTTTAGGTAGAGACTGGGTTTACCATTGTTGGCCAGGCTGGTCTCGAACTCC
 TGACCTCAAATTATCCACCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCC
 ACTGTGCCCAGCCTCCATGTTGTTTTTACAAACACCTGTATCATTTACATTTCCACCAAC
 AGTACACAAGAATTTAGTTTCTCCACATCCTTGCTAGCAGTTGTTATTATCTGTTTTTT

15255 TGGAATCATATGGTATTTGTCTTTTAGTGACTGCCTAATTTAAAAAAATTTTTTTGAGA
 CGGAGTCTCTGCTCTGTGCGCCAGGCTGGAGTGCAGTGGCACCATCTCTGCTCACTGCAAC
 CTCCACCTCCCAGGTTCAAGTGATTGTCTGCTCGGCCTCCACGTAGCTGGGATTACA
 GGTGCTCGCCACAACACCCGGCTAATTTTTGTATTTTTAGGTAGAGACTGGGTTTACCA
 TGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAATTATCCACCTGCCTTGGCCTCCCA
 [A, C]
 AGTGCTGGGATTACAGGCGTGAGCCACTGTGCCCAGCCTCCATGTTGTTTTTACAAACAC
 CTGTATCATTTACATTTCCACCAACAGTACACAAGAATTTAGTTTCTCCACATCCTTGC
 TAGCAGTTGTTATTATCTGTTTTTTTTTAATGGTTTCTTTTTTCTTTTTTCTTTTTTTTT
 TTTTTTTGAGACGGTCTTATTCTTGCTCATGCTGGAGTGCAGTGGTGAATGTGATAGCTC
 ACTGCAGCCTCAACCTCTGGGCTCAAGCAGTCTGCCACCTCAGCCTCCACATAGGTGGG

18996 GTAGTGGCTCATGCCTAAATCTCAGCACTTTGGGAGGCTGAGGCGGGAGGATCACCTGAG
 GTCGGGAGTTTGAGACCAGCCTGGGCAACATGGTGAAACCCCGTGTCTACTAAAAATATA
 AAAACTAGCCAGTCATGGTGGCACACACCTGTAATCCCGGCTACTTGGGAGGCTGAGGCA
 TGAGAATCACTTGAACCTGGGCTGCGGAGGTTGTAGTAAGCTGAGATCATGCCACTGTAC
 TCCAGCCTGGGTGACAGGGAGACATTCTCTCTCTCAAAAAAAAAAAAAAAAAACAAAAAAA
 [A, C]
 AAAACAAACCAACAAAACAAAGTAATCCAGGAACAACAACATGATGAAGGACTGCATGCA
 GGACTCAGTGATGGATGGTGGAAGACAGCCAGGAAGTTAAGCATGACTCTGGTATTAAGT
 GTTGTCTGGGAGAGTTAAGATTCCATTTACAGAAATAAGACCTGTAGGGGAAGCTCTTGA
 TTTTTTTTTTTTTGCAGACTGCTGATTTCTGATTACATGTGTTAAGTTTGAGGTATAGA
 GAGAAAGAACATCCTGGCCGGGTGCAGTGGCTCACACCCGTAATCCCAGCACTTTGGGAG

18996 GTAGTGGCTCATGCCTAAATCTCAGCACTTTGGGAGGCTGAGGCGGGAGGATCACCTGAG
 GTCGGGAGTTTGAGACCAGCCTGGGCAACATGGTGAAACCCCGTGTCTACTAAAAATATA
 AAAACTAGCCAGTCATGGTGGCACACACCTGTAATCCCGGCTACTTGGGAGGCTGAGGCA
 TGAGAATCACTTGAACCTGGGCTGCGGAGGTTGTAGTAAGCTGAGATCATGCCACTGTAC
 TCCAGCCTGGGTGACAGGGAGACATTCTCTCTCTCAAAAAAAAAAAAAAAAAACAAAAAAA
 [A, C]
 AAAACAAACCAACAAAACAAAGTAATCCAGGAACAACAACATGATGAAGGACTGCATGCA
 GGACTCAGTGATGGATGGTGGAAGACAGCCAGGAAGTTAAGCATGACTCTGGTATTAAGT
 GTTGTCTGGGAGAGTTAAGATTCCATTTACAGAAATAAGACCTGTAGGGGAAGCTCTTGA
 TTTTTTTTTTTTTGCAGACTGCTGATTTCTGATTACATGTGTTAAGTTTGAGGTATAGA
 GAGAAAGAACATCCTGGCCGGGTGCAGTGGCTCACACCCGTAATCCCAGCACTTTGGGAG

18996 GTAGTGGCTCATGCCTAAATCTCAGCACTTTGGGAGGCTGAGGCGGGAGGATCACCTGAG
 GTCGGGAGTTTGAGACCAGCCTGGGCAACATGGTGAAACCCCGTGTCTACTAAAAATATA
 AAAACTAGCCAGTCATGGTGGCACACACCTGTAATCCCGGCTACTTGGGAGGCTGAGGCA
 TGAGAATCACTTGAACCTGGGCTGCGGAGGTTGTAGTAAGCTGAGATCATGCCACTGTAC
 TCCAGCCTGGGTGACAGGGAGACATTCTCTCTCTCAAAAAAAAAAAAAAAAAACAAAAAAA
 [A, C]
 AAAACAAACCAACAAAACAAAGTAATCCAGGAACAACAACATGATGAAGGACTGCATGCA
 GGACTCAGTGATGGATGGTGGAAGACAGCCAGGAAGTTAAGCATGACTCTGGTATTAAGT
 GTTGTCTGGGAGAGTTAAGATTCCATTTACAGAAATAAGACCTGTAGGGGAAGCTCTTGA
 TTTTTTTTTTTTTGCAGACTGCTGATTTCTGATTACATGTGTTAAGTTTGAGGTATAGA

FIGURE 3, page 13 of 17

GAGAAAGAACATCCTGGCCGGGTGCAGTGGCTCACACCCGTAATCCCAGCACTTTGGGAG

20532 TAGCCATCATGGAATACTATGCACCCATTAAAAATAATGAAATAATGTCTTTGCAACAAC
ATGAATGTAGCTGGAGGGCATTATCCTAAGCAAATAACACAGAAACAGAAAACCAAATA
CTGCGTGTTCTCACGCAGTGAGAGTGGGAGCTGAACATCAAGTACACATGGATGTAAGA
TGGCAACAATAGACATGGGTCTACTAGAGGTGGTGGTGGCGCAGGGTGGGGGTGGGGGT
GTGTGGCAGAGGAACAGCTGAAAACTACCTATTTGATACTATACCCAGCACCTGGGAAA
[C, T]
GGGTTCACTCATACCCCAAACCTCAGCATCACACAGTATACCTTTCTAACAACCTTACAC
ATGTATTCTGTGATTCTAAAATAAACATTGAAAAATAAAAAAACTGACATGGTTTGT
ACTGTTTAATCTGACATAATGGCTAGGGGAAATGAAGTCTGCAGAATGGCTGTTTACGGA
TGTTGTTGTTGTTGTTGTTGAGATGAGGTCTCACTATGTTGCCAGGCTTGAACCTCTGCCT
CAAGCAGTCTCTGCTGCTTGACCTCCCAAATGTTGAGATTACATGCATGAGCCATTGCC

21118 GCATGAGCCATTGCCAAAACGGCTATTTGGATTGCTGTTAAGGTTATTACATTCTCTGTG
TAGTAAGACCTTGAAGGAGAAGGATTGAGATCAGGAGTTTAAAGAAAAATGTTAATCTA
GGAAGAGAGGATAATTTCTGTGTTTGGCCAGTTGCAATGGCTCACGCCTGTAATCCCAGC
GTTTTGGGAGGCCGAGCTGGGCAGATCACTGAGCTCAGGAGTTTGAGAAGAGCCTGGGC
AACATGGTGAAGCCCCGTCTCTACTGAAAATACAAAATTGGCTGGGCGTGGTGGCAGGT
[G, T]
CCCCTGTAATTTAGCCACTCAGGAGACTGAGGCAGGAGAATTGCTTGAACCTGGAAAG
TGGAGGATATGGTGAGCCTAGATCGGCCACTCTACTCCAGCCTGGGAGACTCCATCTCA
AAAAAAAAGAAAACAAGATGCTGAAATGAAGTAATTACCACAGTCAATGTGATCCTATA
ACTTTGTTTTCTTTAGAGATGGGTCTCCCTCTGTCAACCCAGGCTGGAGTGCAGTGGTG
CATCATAGCTTTCTGCAGCCTCCACCTCCTGGGCTCAGGTGGTCTCTTGCCTCAGTGTT

21211 AGGAGTTTAAAGAAAAATGTTAATCTAGGAAGAGAGGATAATTTCTGTGTTTGGCCAGTT
GCAATGGCTCACGCCTGTAATCCCAGCGTTTGGGAGGCCGAGCTGGGCAGATCACTTGA
GCTCAGGAGTTTGAGAAGAGCCTGGGCAACATGGTGAAGCCCCGTCTCTACTGAAAATAC
AAAAATTGGCTGGGCGTGGTGGCAGGTGCCACTGTAATTTAGCCACTCAGGAGACTGA
GGCAGGAGAATTGCTTGAACCTGGAAAGTGGAGGATATGGTGAGCCTAGATCGCGCCACT
[C, G]
TACTCCAGCCTGGGAGACTCCATCTCAAAAAAAAAGAAAACAAGATGCTGAAATGAAGT
AATTACCACAGTCAATGTGATCCTATAACTTTGTTTTCTTTAGAGATGGGTCTCCCTC
TGTCACCCAGGCTGGAGTGCAGTGGTGCATCATAGCTTTCTGCAGCCTCCACCTCCTGGG
CTCAGGTGGTCTCTTGCTCAGTGTCCGAGTAGTTAGGACTGACTGCAGGTGCATGCT
GCTATGCCTGGCTAACTTTAAATTTTTTTGTAGAGGCGGGTCTTGCTATGTTGCCTAG

21212 GGAGTTTAAAGAAAAATGTTAATCTAGGAAGAGAGGATAATTTCTGTGTTTGGCCAGTTG
CAATGGCTCACGCCTGTAATCCCAGCGTTTGGGAGGCCGAGCTGGGCAGATCACTTGA
CTCAGGAGTTTGAGAAGAGCCTGGGCAACATGGTGAAGCCCCGTCTCTACTGAAAATACA
AAAATTGGCTGGGCGTGGTGGCAGGTGCCACTGTAATTTAGCCACTCAGGAGACTGAG
GCAGGAGAATTGCTTGAACCTGGAAAGTGGAGGATATGGTGAGCCTAGATCGCGCCACTC
[T, C]
ACTCCAGCCTGGGAGACTCCATCTCAAAAAAAAAGAAAACAAGATGCTGAAATGAAGTA
ATTACCACAGTCAATGTGATCCTATAACTTTGTTTTCTTTAGAGATGGGTCTCCCTCT
GTCACCCAGGCTGGAGTGCAGTGGTGCATCATAGCTTTCTGCAGCCTCCACCTCCTGGGC
TCAGGTGGTCTCTTGCTCAGTGTCCGAGTAGTTAGGACTGACTGCAGGTGCATGCTG
CTATGCCTGGCTAACTTTAAATTTTTTTGTAGAGGCGGGTCTTGCTATGTTGCCTAGG

21238 GGAAGAGAGGATAATTTCTGTGTTTGGCCAGTTGCAATGGCTCACGCCTGTAATCCCAGC
GTTTTGGGAGGCCGAGCTGGGCAGATCACTGAGCTCAGGAGTTTGAGAAGAGCCTGGGC
AACATGGTGAAGCCCCGTCTCTACTGAAAATACAAAATTGGCTGGGCGTGGTGGCAGGT
GCCACTGTAATTTAGCCACTCAGGAGACTGAGGCAGGAGAATTGCTTGAACCTGGAAA
GTGGAGGATATGGTGAGCCTAGATCGCGCCACTCTACTCCAGCCTGGGAGACTCCATCTC
[-, A]
AAAAAAAAGAAAACAAGATGCTGAAATGAAGTAATTACCACAGTCAATGTGATCCTATA
ACTTTGTTTTCTTTAGAGATGGGTCTCCCTCTGTCAACCCAGGCTGGAGTGCAGTGGTG
CATCATAGCTTTCTGCAGCCTCCACCTCCTGGGCTCAGGTGGTCTCTTGCCTCAGTGTT
CCGAGTAGTTAGGACTGACTGCAGGTGCATGCTGCTATGCCTGGCTAACTTTAAATTTTT
TTGTAGAGGCGGGTCTTGCTATGTTGCCTAGGCTGGTCTCAACTCCTGATCTCAATC

26483 CTGTAAGCATTGTATATACATTATCTTTGTGATTGTGATGGGGTCTTCTGGTTCCTGC

FIGURE 3, page 14 of 17

TAGTATTATGTGCTTTTTTTTCCCCTCAAGACTGGAGCAGTTATTAGCCCCAATAGCCA
 ATCATTAAAGCCTAAATCCTAATTCACAGTAGCATTGTGGGCTTCCTGGATCCTCAGCCAG
 AATAGGGTTTTTACAACCTAACAAATAAAAAATGAGACGTCAGAGGGGAAGTATAGTAACT
 AGTGTGTTTTTGATTAAGAAGGGGATGAAACACAAAAACCAAAGAAGTCTGTGGAGGAG
 [G,A]
 AGGAGCTAGGGCATGTTCTTCTGAGACTTGAGCGAGAGGAACCTTGGGAGTGGGAGGTTG
 TGGGGAAGTTAGAGGCTGCAAGGGCTGTGAGGTAGTGAGAGGGACGGATCCCATGAGGA
 GTCTGGCATGGGGGCTCTGATTTAGCCTCTTCCCTGCAGTGGACAATGAGGAGGCTGCAC
 TGCTGCATGAAGAGGCTACCATGACTATTGAAGAGCTGCTGACACGCTACGGGCAGAAT
 GTCACAAGGGCCCTCCCACAGCAAATCTGGAGGTGGGACAGGCGAGGAACCAGGGTCCC

GENEWISE ALIGNMENT:

| | | | |
|----------------|-------|--|-------------------------|
| gi 4505999 ref | 1 | MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPEL | |
| | | MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSM | |
| | | MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSM | |
| genomic | 2134 | aggtctccaagattggggggccccctgttgacgtcgttaggaacgcagt | |
| | | tgcatacactagcgagtgccgtctcagtcctagggctatggagctgg | |
| | | gtccccgccgggcccgcggcgccccgggcccccgacgccccgggggggcacg | |
| gi 4505999 ref | 50 | DSETAMFSVYDGHGG | EEVALYCA |
| | | ++ + G | EEVALYCA |
| | | PLRGGNLTEKESGGW | EEVALYCA |
| genomic | 2281 | cccgacagagaggtGGTCCTCC Intron 1 TAGGGgggggtttg | |
| | | ctgggatcaaagggg <1-----[2327:25451]-1> aatctagc | |
| | | gggcgtgggagcgag | gacgctc |
| gi 4505999 ref | 73 | KYLPDIKDQKAYKEGKLQK | ALEDAF |
| | | KYLPDIKDQKAYKEGKLQK | ALEDAF |
| | | KYLPDIKDQKAYKEGKLQK | ALEDAF |
| genomic | 25478 | atccgaaagcagtaggaccaGTCTGTC Intron 2 CAGgtgggt | |
| | | aatcattaaaacaaagataa<0-----[25538:25851]-0>ctaact | |
| | | attttccatggccgacgagg | taatcc |
| gi 4505999 ref | 99 | LAIKALKTTEEVIKELAQIAGRPTEDDEKEKVADEDD | |
| | | LAIKALKTTEEVIKELAQIAGRPTEDDEKEKVADEDD | |
| | | LAIKALKTTEEVIKELAQIAGRPTEDDEKEKVADEDD | |
| genomic | 25870 | tgaggataaggggaagcgagccagggggagagggggg | |
| | | tctacatccaattaatcatcgcccaaaaaaatcaaaa | |
| | | gttccagctaactaggagtagactgtataaaaattatt | |
| gi 4505999 ref | 137 | | DNEEAALLHEEATMTIEELLTRY |
| | | | DNEEAALLHEEATMTIEELLTRY |
| | | V:V[gtg] | DNEEAALLHEEATMTIEELLTRY |
| genomic | 25984 | GGTGAGTG Intron 3 CAGTGgagggggcccggaagggccact | |
| | | <1-----[25985:26642]-1> aaaaccttaacctctaattcga | |
| | | | ctggtaggtagtcgttagggacc |
| gi 4505999 ref | 161 | GQNCHKGPPhSKSGGGTGEEPGSQGLNGEAGPEDSTRETSPQENGPTAK | |
| | | GQNCHKGPPhSKSGGGTGEEPGSQGLNGEAGPEDSTRETSPQENGPTAK | |
| | | GQNCHKGPPhSKSGGGTGEEPGSQGLNGEAGPEDSTRETSPQENGPTAK | |
| genomic | 26714 | gcatcagcccaatgggagggcgctcgagggcggttaagactcgagcaga | |
| | | gaagaagccagacgggcgaacgcagtagacgcaaccgacccaaagccca | |
| | | ggctcgctcccatatgacgaagcgccctggaatgcatgattaaatccacg | |
| gi 4505999 ref | 210 | AYTGfSSNSERgTEAGQVGEPIPTGEAGPSCSSASDKLPRVAKSKFFE | |
| | | AYTGfSSNSERgTEAGQVGEPIPTGEAGPSCSSASDKLPRVAKSKFFE | |
| | | AYTGfSSNSERgTEAGQVGEPIPTGEAGPSCSSASDKLPRVAKSKFFE | |
| genomic | 26861 | gtagtttatgagggggcgacaggggcttttgtgacccggatattg | |
| | | cacgtccacaggcacgatgacgtccgacgcccgaatcgtcacatta | |
| | | ccactcccgatgtgacattgtctcttgtgtcctactcggtattgagcgtg | |
| gi 4505999 ref | 259 | DSEDESDAEEDSE | ECSEEDGY |

FIGURE 3, page 15 of 17

| | | | |
|----------------|-------|---|------------------------------|
| | | DSEDESDEAEEDSE | ECSEEDGY |
| | | DSEDESDEAEEDSE | ECSEEDGY |
| genomic | 27008 | gaggggtgggggggagGTAAGGG Intron 4 | CAGgtaggggt |
| | | agaaacaacaaaaaga<0-----[27059:27638]-0>aggaaaaa | |
| | | ctgtgatggaagactg | accgagtcc |
| gi 4505999 ref | 285 | SSEEAENEDEDDTEEAEDDEEEEMMVPMEGKEE | |
| | | SSEEAENEDEDDTEEAEDDEEEEMMVPMEGKEE | |
| | | SSEEAENEDEDDTEEAEDDEEEEMMVPMEGKEE | |
| genomic | 27666 | aaggggagggggagggggggggggaagcgaggagg | |
| | | ggaacaaaaaaacaacaaaaaaatttcgtagaaa | |
| | | ctggagtgatgtccggtagctaaaaagggaggacagg | |
| gi 4505999 ref | 323 | | PGSDSGTTAVVALIRGKQLIVANAGD |
| | | | PGSDSGTTAVVALIRGKQLIVANAGD |
| | | | PGSDSGTTAVVALIRGKQLIVANAGD |
| genomic | 27780 | GTGTGTG Intron 5 | AAGcgtgagaaggggcacgactaggagg |
| | | <0-----[27780:28130]-0>cgcaggcccttcttggaatttcacga | |
| | | | tctcttaagggcgaaggggtaccaac |
| gi 4505999 ref | 349 | SRCVVSEAGKALDMSYDHKPEDEVELARIKNAGGKVTMDGRVNGGLNLS | |
| | | SRCVVSEAGKALDMSYDHKPEDEVELARIKNAGGKVTMDGRVNGGLNLS | |
| | | SRCVVSEAGKALDMSYDHKPEDEVELARIKNAGGKVTMDGRVNGGLNLS | |
| genomic | 28209 | tctggtgggagtgattgcacggggcgcaaagggagaagcgaggcact | |
| | | cggttcacgactatcaaaacaaatcgtaacggatctaggttaggtatc | |
| | | tctgatgtcatagcttcaagtaaaaaaccgtttcgccgtgaccgcccc | |
| gi 4505999 ref | 398 | RAI | DHFYKRKNLNPPEEQMISAL |
| | | RAI | DHFYKRKNLNPPEEQMISAL |
| | | RAI G:G[ggg] | DHFYKRKNLNPPEEQMISAL |
| genomic | 28356 | agaGGTAAGGG Intron 6 | CAGGGgcttaaaaaccgggcaatgc |
| | | gct <1-----[28366:29164]-1> | aataagaaatccaaattcct |
| | | act | ccctgacgcgatgaggtact |
| gi 4505999 ref | 422 | PDIKVLTLTDDHEFMVIACDGI | N |
| | | PDIKVLTLTDDHEFMVIACDGI | N |
| | | PDIKVLTLTDDHEFMVIACDGI W:W[tgg] | N |
| genomic | 29227 | cgaagcacaggcgtagagtgggTGGTGAGCA Intron 7 | CAGGa |
| | | catattctcaaaattttcgagt <2-----[29295:29541]-2> | a |
| | | tccgggtctcctacgctcttcc | t |
| gi 4505999 ref | 446 | VMSSQEVVDFIQSKISQRDENGELRLLSSIVEEL | |
| | | VMSSQEVVDFIQSKISQRDENGELRLLSSIVEE+ | |
| | | VMSSQEVVDFIQSKISQRDENGELRLLSSIVEEV | |
| genomic | 29546 | gaaacgggggtactaaaccggaggccttttagggg | |
| | | ttggaattattacatgagaaagatgttccttaat | |
| | | ggccgatatctaagccgttatggtgagactgagg | |
| cDNA | 1 | MGAYLSQPNTVKCSGDVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPEL | |
| | | MGAYLSQPNTVKCSGDVGAPRLPLPYGFSAMQGWRVSME | |
| | | MGAYLSQPNTVKCSGDVGAPRLPLPYGFSAMQGWRVSMVRRQGPIGW | |
| genomic | 2134 | aggtctccaagattggggggccccctgttgacgtcgttaggaacgcagt | |
| | | tgcatcacactagcgagtgccgtctcagtcctaggtctatggagctgg | |
| | | gtccccgccgggcccgcgcccgcgggcccccgacgcccgggggggacg | |
| cDNA | 50 | DSETAMFSVYDGHGG | EEVALYCA |
| | | ++ + G | EEVALYCA |
| | | PLRGGNLTEKESGGW ~:G[ggg] | EEVALYCA |
| genomic | 2281 | cccggacagagaggtGGTCCTCC Intron 1 | TAGGGgggggttg |
| | | ctgggatcaaagggg <1-----[2327:25451]-1> | aatctagc |
| | | gggcgtgggagcgag | gacgctc |

FIGURE 3, page 16 of 17

| | | | | |
|---------|-------|--|-------------------------|------------------------------|
| cDNA | 73 | KYLPDI IKDQKAYKEGKLQK | | ALEDAF |
| | | KYLPDI IKDQKAYKEGKLQK | | ALEDAF |
| | | KYLPDI IKDQKAYKEGKLQK | | ALEDAF |
| genomic | 25478 | atccgaaagcagtaggaccaGTCTGTC | Intron 2 | CAGgtgggt |
| | | aatcattaaaaacaaagataa<0-----[25538:25851]-0> | | ctaact |
| | | atttccatggccgacgagg | | taatcc |
| cDNA | 99 | LAI DAKLTTEEVIKELAQIAGRPTED EDEKEKVADEDD | | |
| | | LAI DAKLTTEEVIKELAQIAGRPTED EDEKEKVADEDD | | |
| | | LAI DAKLTTEEVIKELAQIAGRPTED EDEKEKVADEDD | | |
| genomic | 25870 | tgaggataaggggaagcgcaggccagggggagagggggg | | |
| | | tctacatccaattaatcatcggccaaaaaaatcaaaa | | |
| | | gttcagctaactaggagtagactgtataaaaattatt | | |
| cDNA | 137 | | | DH FYKRKNLPP EEQMISALPDI |
| | | | | DH FYKRKNLPP EEQMISALPDI |
| | | G:G[ggg] | | DH FYKRKNLPP EEQMISALPDI |
| genomic | 25984 | GGTGAGTG | Intron 3 | CAGGGgcttaaaaacccggcaatgccga |
| | | <1-----[25985:29164]-1> | | aataagaaatccaaattcctcat |
| | | | | ccctgacgcgatgaggtacttcc |
| cDNA | 161 | KVLTLTD DHEFMVIACDGI | | NVMS |
| | | KVLTLTD DHEFMVIACDGI | | NVMS |
| | | KVLTLTD DHEFMVIACDGI | W:W[tgg] | NVMS |
| genomic | 29236 | agcacaggcgtagagtggatGGTGAGCA | Intron 4 | CAGGagaa |
| | | attctcaaaattttcgagt | <2-----[29295:29541]-2> | attg |
| | | gggtctcctacgtcttcc | | tggc |
| cDNA | 185 | SQEVVDFIQSKISQRDENGELRLLS SIVEEL | | |
| | | SQEVVDFIQSKISQRDENGELRLLS SIVEE+ | | |
| | | SQEVVDFIQSKISQRDENGELRLLS SIVEEV | | |
| genomic | 29555 | acgggggtactaaaccggaggccttttagggg | | |
| | | gaattattacatgagaaagatgttccttaat | | |
| | | cgatatctaagccgttatggtgagactgagg | | |

FIGURE 3, page 17 of 17